UNIVERSITY OF LONDON THESIS

Degree PhD Year 2007 Name of Author WOODMAN, Nicholas Daniel

MODELLING OF TRANSPORT IN HIGHLY HETEROGENEOUS POROUS MEDIA, WITH APPLICATION TO THE FLUSHING OF WASTE

Nicholas Daniel Woodman

UNIVERSITY COLLEGE LONDON

2007

Submitted in fulfilment of the degree of Doctor of Philosophy

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Authorship

I, ***Nicholas Daniel Woodman***, confirm that the work presented in this thesis is my own. Where information that has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis is concerned with predicting the fate of contaminants using tracer and contaminant flushing datasets in otherwise poorly characterised heterogeneous media.

A deductive approach towards constraining prediction uncertainty is adopted, by narrowing a 'pool' of plausible process models and associated parameter sets. Mathematical comparison of simple transport models is used to allow enhanced diagnosis. Transport models are put into a framework. Then they are compared using commonly used measures of breakthrough curves (least squares fitting, method of moments and late-time concentration gradients).

Experimental data and models for transport in waste are then reviewed using this framework. Two new tracer datasets, at laboratory and lysimeter scales, are analysed and modelled. Strong evidence is shown of 'dual-porosity' diffusive exchange effects being an important component in transport in municipal solid waste. However, it is not possible to discriminate between contributions to dispersion through advective and diffusive effects. Approaches to enhance the diagnostic capability of future waste experiments are developed.

One possible 'entrant' to the 'pool' of models is examined: a stochastic channel network model. Histograms of nodes and average heads are evaluated at different distances from a range of boundary conditions for large ensembles revealing key insights into network flow behaviour near different types of boundary. Key properties of the network are related to typical continuum analyses, including dimensionality of flow and genesis of apparent 'skins' near boundaries. 'Classical' percolation networks (with zero spatial correlation) are also analysed, revealing new insights into their network properties within finite boundaries.

In summary, progress towards predicting the fate of contaminants in highly heterogeneous systems is made through a thorough delineation of the behaviour of simple models in tandem with adoption of a fertile approach to adding new conceptual models. The increased rigorousness of using multiple diagnostic criteria in conjunction with adopting a philosophy of working with multiple working hypotheses is commended as a methodology that provides a sound basis from which to employ multiple-model uncertainty analysis for predictive purposes in the future.

Acknowledgements

My first thanks must go to my primary supervisor, John Barker, whose influence on me personally has been incalculable. I feel very lucky to have had the chance to learn from John, benefiting from both his kindness and wealth of ideas. Although it scarcely needs doing, I hope this thesis highlights both the breadth and importance of many of his papers in the literature and points to their continuing relevance.

It has also been a pleasure to have been supervised by Richard Beaven, both on a personal level and also because of his wide knowledge of hydrogeology, engineering and waste science and his constant help and support.

This thesis relies on a large body of data, for which I am indebted to Andrew Hudson, Richard Beaven, Layi Oni, Julia Stegemann and Jing Lin, who are all warmly thanked.

I have also been lucky enough to have worked for John Black during the PhD period, examining fundamentals of channel-network behaviour. For this work a model, written by Peter Robinson of Quintessa, was extremely valuable.

Dr Noelle Odling and Dr Adrian Butler are thanked for their constructive criticism, particularly in encouraging more clarity in demonstrating the application of the methodology developed in Part I of the thesis to the data in Part II.

Thanks to Nic Higgins for kindly proof-reading the final manuscript.

Thanks to my family who have put up with absences and distractedness. Particular thanks are due to John and Caroline, particularly in the grandson department. Thanks to Jo, who has been my utter support throughout.

The EPSRC are thanked for funding this research.

Dedication

To Dr Roger Morris,
who is greatly missed.

Contents

Tables

Figures

Notation

Roman symbols

Symbol	Definition	Dimension
$\underline{a}$	Parameter vector.	[Depends on parameter]
a	Fracture width .	[L]
A	Cross sectional area.	[L^2]
A_n	Coefficient in infinite sum solutions relating to slab-geometry. $A_n = (2n-1)^2\pi^2/4$	[-]
b	Characteristic block dimension, block volume/block area.	[L]
B	Bias.	[Depends on parameter]
$\mathrm{B}(\sqrt{s})$	Block Geometry Function, $\mathrm{B}(\sqrt{s}) = \dfrac{\langle \overline{C}_{im} \rangle}{\overline{C}_m}$.	[-]
$\mathrm{B}_t(t)$	Block Geometry Function in time domain.	[-]
$\hat{\mathrm{B}}^{(n)}$	$\hat{\mathrm{B}}^{(n)}(0,t_{cb}) = \dfrac{\partial^n}{\partial s^n}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big\vert_{s=0}$	[T^n]
$\hat{\mathrm{B}}'$	$\hat{\mathrm{B}}'(0,t_{cb}) = \dfrac{\partial}{\partial s}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big\vert_{s=0}$	[T]
B'(x)	$\mathrm{B}'(x) = \dfrac{d}{dx}\mathrm{B}(x)$	[-]
B_d	Bulk density.	[ML^{-3}]
B_n	Coefficient in infinite sum solutions relating to spherical-geometry. $B_n = n^2\pi^2$	[-]
c	Local (microscopic) concentration.	[ML^{-3}]
C	Concentration. $C = C^r$ unless otherwise stated.	[ML^{-3}]
C^f	'Flux-averaged' concentration.	[ML^{-3}]
C_{im}	Concentration in immobile zone.	[ML^{-3}]
C_m	Concentration in mobile zone.	[ML^{-3}]
C_o	Input concentration.	[ML^{-3}]
C^r	'Resident' or volume-averaged concentration.	[ML^{-3}]

Symbol	Definition	Dimension
C_{Real}	The concentration generated by a model which is assumed to be 'reality' before noise is added in order to make synthetic data.	[ML^{-3}]
C_δ , C_H	Impulse and step inputs of concentration respectively.	[ML^{-3}]
C_B	Concentration before noise added.	[ML^{-3}]
C_I	Initial concentration in the system	[ML^{-3}]
C_S	Specific conductance. Conductance per unit length (or conductivity times area). Note that $K_{\text{sub-channel}} = C_S/\Delta L^2$	[L^3T^{-1}]
C^*	Sorbed mass per unit dry solid.	[-]
$[C_{ij}]$	Covariance matrix.	[Depends on parameter]
D^*	Coefficient of molecular diffusion	[L^2T^{-1}]
D	(Hydrodynamic) Longitudinal Dispersion coefficient. $D = \alpha V$, or $D = \alpha V + D_a$	[L^2T^{-1}]
D_a	Apparent diffusion coefficient.	[L^2T^{-1}]
D_E	Effective or intrinsic diffusion coefficient.	[L^2T^{-1}]
D_F	Flow dimension. The apparent dimensionality assuming flow is through a uniform isotropic continuum.	[-]
D_M	Macrodispersion coefficient, $D_M = \dfrac{1}{2}\dfrac{\partial\sigma_z^2}{\partial t} = \dfrac{V}{2}\dfrac{\partial\sigma_z^2}{\partial z}$	[L]
D_T	Free water diffusion coefficient	[L^2T^{-1}]
D_{Tay}	Dispersion coefficient evaluated using Taylor's solution.	[L^2T^{-1}]
e , E	Error.	[Depends on parameter]
$E[x]$	Expectation of x (mean).	[Depends on parameter]
E_i	Expected (modelled) value	[Depends on parameter]

Symbol	Definition	Dimension
E_N	Nash-Sutcliffe efficiency, $E_N = 1 - SSE / SS_{yy}$ $= 1 - MSE / \sigma_O^2$	[-]
F, f	General function.	
$f(C)$	Sorption isotherm.	[-]
g(t)	$g(t) = \sigma \mathrm{B}(t)$	[-]
$\overline{G}$	Mean Gap length (i.e. mean distance between channels).	[sub-channels]
h	Slab half-thickness.	
h	Head.	[L]
h_S	Skin. An apparent head that is the offset between the predicted head at the cylindrical boundary and the true value when head is projected by a linear regression.	[L]
h_{LS}	Linear Skin. An apparent head that is the offset between the predicted head at the F2F boundary and the true value when head is projected by a linear regression.	[L]
H(t)	Heaviside step function.	
HW	Half-width of a uniform distribution.	[Depends on parameter]
i	Index variable.	[-]
I_γ	Modified Bessel function of the first kind of order γ.	[-]
K	Hydraulic conductivity.	[LT^{-1}]
K_M	Young and Ball's dimensionless MIM transfer coefficient, $K_M = \frac{\kappa t_{cb}}{\sigma t_a}$.	[-]
K_d	Distribution coefficient (linear sorption).	[L^{3}M^{-1}]
K_{eff}	Effective permeability. The bulk permeability for a fully occupied network in which the sub-channels had lognormally-distributed conductance.	[LT^{-1}]
K_G	Geometric mean hydraulic conductivity.	[LT^{-1}]
K_h	Horizontal hydraulic conductivity	[LT^{-1}]
K_v	Vertical hydraulic conductivity	[LT^{-1}]
L , L^{-1}	Laplace transform , inverse Laplace transform.	[-]
L	Channel length.	[sub-channels]
L	Distance to end of system.	[L]
L	Unit of Litres	[L^3]
$\overline{L}$	Mean channel length. The average channel length out of all the channels in a realisation.	[L]
m	Heterogeneity constant for the Langmuir-Freundlich isotherm.	[-]
m_0	Injected mass per unit area.	[ML^{-2}]
M	Number of parameters.	[-]
M_0	Total mass injected.	[M]
n, N	Count number.	[-]
n	Ensemble size. Number of realisations in an ensemble.	[-]
N	Freundlich exponent	[-]
$\mathrm{N}(B, \sigma_N^2)$	Normal distribution.	[-]
N_A	Number of active nodes.	[-]
N_B	Number of branch nodes.	[-]
N_F	Number of flowing nodes.	[-]
N_P	Number of pore volumes.	[-]
O	Observed value.	[Depends on parameter]
$\hat{P}_{ON}^{1D}$	Node density (1D). Probability that a node is adjacent to a sub-channel (along a line).	[-]
$\hat{P}_{ON}^{3D}$	Node density (3D). Probability that a node is adjacent to a sub-channel.	[-]
P_P	Percolation probability. Probability that a given bond belongs to the boundary-spanning cluster, given a fraction of occupied bonds, P.	[-]
P_C	Percolation threshold. Probability above which 50% of realisations percolate.	[-]

Symbol	Definition	Dimension
P	Probability.	[-]
P	Bond occupancy. Number of possible bond (or sub-channels) which have been allocated divided by total number of possible bonds in the network.	[-]
P_P	Proportion percolating. Number of realisations (runs) which percolate per total number of runs in the ensemble.	[-]
P_A	Probability of a another sub-channel following an existing sub-channel along a line.	[-]
P_N	Probability of a new channel starting if there was no channel previously.	[-]
P_{ON}	Channel density. Probability that a sub-channel is active along a line.	[-]
Pe	Peclet Number, $Pe = \frac{Vz}{D_a}$.	[-]
Pe_g	Grain Peclet number (d is the characteristic grain dimension), $Pe_g = \frac{Vd}{D_a}$.	[-]
q	Water flux rate, or specific discharge (also known as Darcy velocity), $q = Q/A$.	[LT^{-1}]
q_e	Exchange term, $q_e = \sigma\Gamma(t)$.	[ML^{-3}T^{-1}]
q_s	Solute flux rate	[ML^{-2}T^{-1}]
Q	Flow rate.	[L^{3}T^{-1}]
r	Radial distance.	[L]
r^2	Correlation coefficient.	[-]
R_D	Ratio of diffusion coefficients, $R_D = \frac{D_{A2}}{D_{A1}}$.	[-]
R	Retardation coefficient.	[-]
s	Standard deviation of $\log(C_S)$.	[L^{3}T^{-1}]
s	Laplace variable.	[T^{-1}]
S	Dimensionless Laplace variable, $S = st_{cb}$	[-]

Symbol	Definition	Dimension
s, SD	Standard deviation.	[Depends on parameter]
t	Time.	[T]
t_a	Mobile zone advection time, $t_a = \frac{z}{V_m} = \frac{Vol_m}{Q}$.	[T]
t_{cb}	Characteristic block diffusion time, $t_{cb} = \frac{b^2}{D_a}$.	[T]
t_{cf}	Characteristic fracture diffusion time, $t_{cf} = \frac{a^2 D_a}{4D_F^2}$.	[T]
t_d	Characteristic hydraulic dispersion time for mobile zone, $t_d = \frac{D}{V_m^2} = \frac{\alpha}{V_m}$.	[T]
t_A	Whole porosity advection time, $t_A = \frac{z}{V} = \frac{Vol}{Q}$.	[T]
t_D	Characteristic hydraulic dispersion time for whole porosity, $t_D = \frac{D}{V^2} = \frac{\alpha}{V}$.	[T]
t_{IM}	Characteristic MIM immobile block time, $t_{IM} = \frac{\theta_{im}}{\kappa}$.	[T]
t_M	Characteristic MIM mobile block time, $t_M = \frac{\theta_m}{\kappa}$.	[T]
T	Dimensionless time, t/t_{cb}.	[-]
T_p	Injection pulse duration	[T]
T_s	Time at which a pause (stop) in the flow commences in a column test.	[T]
V	Linear velocity for flow through the total porosity, $V = \frac{q}{\theta}$.	[LT^{-1}]
V_m	Linear velocity in mobile zone,	[LT^{-1}]

Symbol	Definition	Dimension
	$V_m = \frac{q}{\theta_m}$.	
V_S	Shock velocity.	[LT^{-1}]
V, Vol	Volume.	[L^3]
w	Weight of a variable (e.g. in a weighted sum).	[-]
x	Argument of the BGF, $x = \sqrt{st_{cb}}$.	[-]
z	Distance.	[L]
Z	Dimensionless distance, $Z = z / L$.	[-]

Greek symbols

Symbol	Definition	Dimension [ ]
α	Dispersivity.	[L]
α	Constant for non-linear sorption. $\alpha = \frac{B_d K}{\theta}$.	[-]
α	Langmuir adsorption constant.	[L^3M^{-1}]
α_M	Macrodispersivity	[L]
$[\alpha]$	Hessian matrix.	[Depends on parameter]
β	Maximum amount of solute than can be adsorped.	[-]
β	$\beta = \frac{\theta_m}{\theta}$	[-]
γ	Scale parameter for a gamma distribution.	[-]
γ	Dimensionless advection time, $\gamma = t_a / t_{cb}$.	[-]
$\Gamma(t)$	Exchange term. $\Gamma(t) = \frac{\partial \langle C_{im} \rangle}{\partial t}$.	[$ML^{-3}T^{-1}$]
$\Gamma(a)$	Gamma distribution.	[-]
δ	Parameter used for numerical dispersion in DPD-PULSE.	[-]
$\delta(t)$	Dirac of t (temporal impulse function).	[T^{-1}]
ΔL	Sub-channel length (distance between two nodes).	[L]
η	Shape parameter for a gamma distribution.	[-]
θ	Total porosity.	[-]
θ_{eff}	Effective porosity.	[-]
θ_m	Mobile porosity.	[-]
θ_{im}	Immobile porosity.	[-]
θ_d	Drainable porosity (porosity taken up by water that will drain freely under gravity).	[-]
κ	First-order (MIM) exchange coefficient. $\theta_{im} \frac{\partial C_{im}}{\partial t} = \kappa (C_m - \overline{C}_{im})$.	[T^{-1}]
κ'	Dimensionless first-order exchange coefficient, $\kappa' = \frac{\kappa t_{cb}}{\theta_{im}}$.	[-]
λ	Mixing time constant. $\lambda = Q / Vol$.	[T^{-1}]
λ	Integral scale, $\lambda = \frac{1}{\sigma^2} \int_0^\infty R(z) dz$, where $R(z)$ is the covariance function and σ^2 is the variance.	[L]
μ_N'	Nth raw moment, $\mu_N' = \int_0^\infty t^N f(t) dt$.	[Depends on parameter in question]
μ_N	Nth central moment, $\mu_N = \int_0^\infty (t - \mu_1')^N f(t) dt$	[Depends on parameter in question]
ρ_D	Dry density. Dry mass per unit volume of medium.	[ML^{-3}]
ρ_W	Wet density. Wet mass per unit volume of medium.	[ML^{-3}]
ρ_L	Leachate density.	[ML^{-3}]
σ	Ratio of immobile to mobile porosity. $\sigma = \frac{\theta_{im}}{\theta_m}$.	[-]
σ	Standard deviation.	[unit of parameter which the SD applies to]
σ_t^2	Second centred moment of time.	[T^2]
σ_N	SD of noise.	[unit of property which the SD applies to]
σ_Y^2	Variance of $Y = \ln(K)$.	[-]
σ_z^2	Second centred moment of space.	[L^2]
τ	Time, dummy variable (for convolution).	[T]
τ	Dimensionless time, $\tau = t_a \sigma N_P / (t_{cb} R)$.	[-]
χ^2	Chi-squared.	[-]

Glossary

Abbreviations

Abbreviation	Definition
AD	Advection-Dispersion.
AD-DP	Advection-Dispersion and Dual-Porosity.
ADE	Advection-Dispersion Equation. Also known as 'CDE' (convection dispersion equation) .
AN	Analytical solution.
APC	Air Pollution Control residues.
ASR	Aquifer Storage and Recovery.
BGF	Block Geometry Function. Introduced by Barker (1985). The definition preferred for use in this thesis is, "impulse response of the immobile zone".
BTC	BreakThrough Curve. Concentration against time of a tracer arriving at a certain place.
CDF	Cumulative Distribution Function (integral of pdf).
CI	Confidence Interval. Confidence range (say 95%) projected for a single parameter.
CLT	Central Limit Theorem.
CR	Confidence Region.
DK	Dual permeability.
DP	Dual Porosity.
DP(slab) / DP(sphere) etc.	Dual Porosity with a Fickian exchange term. Block geometry is given in brackets
DP-MIM	Dual Porosity with a MIM exchange term.
EC	Electroconductivity.
EPM	Equivalent Porous Medium.
EDZ	Excavation Damage Zone. Zone of supposedly altered permeability around a drift due to the disturbance caused by construction.
F2F	Face to Face. A system that is flowing between two opposite faces on a cube.
FD	Finite Difference.
FE	Finite Element.
GLUE	Generalised Likelihood Uncertainty Estimation.
GOF	Goodness-Of-Fit.
ID	Internal Diameter.
iid	Independent and identically distributed.
LS	Liquid to Solid ratio
LS	Least squares.
LT	Laplace transform.
MC	Monte-Carlo.
ME	Mean Error.
MIM	Mobile-Immobile. Also known as TRM (two-region model), dead-end pore model, PSS pseudo-steady-state,quasi-steady state QSS, 1st-Order model, LDF Linear driving force.
MoI	Moment of Inertia.
MOM	Method Of Moments.
MSE	Mean square error, $\frac{1}{N}\sum_{i=1}^{N}(O_i - E_i)^2$.
MSW	Municipal Solid Waste.
MSWI	Municipal Solid Waste Incineration residue.
NMC	Number of Monte-Carlo realisations.
ODE	Ordinary Differential Equation.
PDE	Partial Differential Equation.
Pdf	Probability Density Function.
RTD	Residence Time Distribution.
SA	Semi-Analytical solution.
SA	Stochastic-Advective.
SADF	Site Affinity Distribution Function (after Kinniburgh et al. (1983)).
SSE	Sum of Square Errors, $\sum_{i=1}^{N}(O_i - E_i)^2$.
SSR	Sum of Square Residuals, $\sum_{i=1}^{N}(E_i - \overline{O})^2$.
SS_{yy}	$\sum_{i=1}^{N}(O_i - \overline{O})^2$.
TDS	Total Dissolved Solids.
TF	Transfer Function.
TOC	Total Organic Carbon.
USS	UnSteady State.
w.r.t	With respect to.

Definitions

Term	Definition
Active node	Node that is connected to both boundaries.
Background concentration	Concentration of traced substance before adding additional pulse.
Bayesian approach	Refinement of prior parameter pdf by confrontation with data to yield a posterior pdf by using Bayes Theorem.
Behavioural model	Beven (2002): a model which gives an 'acceptable' fit to the data.
Bond	A connection between two sites (or nodes). This is the same as a sub-channel.
Branch node	Node that is connected to both boundaries (i.e. active), but has no flow.
Calibration	Varying of parameter values until the (comparable) model output achieves an acceptable minimal difference to data. Synonymous with fitting.
Channel	A series of connected sub-channels in a line.
Choke	A sub-channel across which there is a relatively large head drop. Synonymous with 'waterfall'.
Classical percolation network	Network for which there is a random allocation of bonds between adjacent nodes.
Conceptual model	Konikow and Bredehoeft (1992): A hypothesis for how a system or process operates.
Conservative	Not reacting or sorbing.
Detection limit	Lowest concentration that can be measured above background.
Determinand	Measured property in the form of basic element, ion or compound or proxy (e.g. EC).
Empirical	Evidence-based rather than based on first principles.
End member	An extreme case (perhaps in terms of simplicity, but also potentially in terms of operation of a particular process).
Ensemble	Collection of results.
Equidimensional	Applied to a finite plane that represents a fracture, which has identical length scales for its two principal axes (e.g. a circle or a square).
Equifinality	Beven (2002): multiple model (and parameter sets) that all produce acceptable simulations.
Flowing node	A node which has at least one channel connected to it that has a non-zero flow through it. Synonymous with 'back-bone'.
Homogenisation	Establishing an 'average' property through volumetric averaging.
Hyper-Convergence	The concept of extreme concentration of flow towards a small number of drift inlets, causing the flow dimension to deviate significantly from cylindrical flow.
Linear skin	The head predicted at the boundary if a linear regression is made for internal heads and distances in a F2F flow system.
Lin-log	Linear-log axes. x linear, y logged.
Locally stochastic	Representation of heterogeneity that adds up the contributions of lumped models.
Log-log	Log-log axes. x and y are logged.
Non-equidimensional	A finite plane that represents a fracture, which has different length scales for its two principal axes (e.g. an ellipse or a rectangle).
Non-identifiability	Beven (2002): Poorly defined (parameter) optimum
Non-uniqueness	Multiple local (parameter) optima.
Objective function	A target function that combines the 'difference' or 'distance' between all the measured and estimated data. The function normally needs to be minimised.

Term	Definition
	Synonymous with 'goal' or 'target' function.
Perceptual model	Beven (2002): a qualitative understanding (not formalised into a mathematical model).
Pool	Set of models which are to be considered (multiple working hypotheses).
Pool	A collection of nodes for which there is no head drop greater than Δh between nodes. The edge of the pool is defined either by the domain boundary or by channels that do have a head drop that exceeds Δh.
Realisation	A single run from the model.
Site	Node.
Stream-tube	A closed volume bounded by streamlines and inlet and outlet surfaces.
Sub-channel	A linkage between two nodes (a single 'bond').
Synthetic data	Model plus noise-model created 'data'. Synonymous with 'pseudo' data.
Top-Hat Pulse	Input as a rectilinear pulse.
Tracer identifiability	Ability to notice the tracer from other substances .
Transfer Function	The ratio of the LT output function to the LT of the input function, $G = \frac{\bar{y}_{OUT}}{\bar{y}_{IN}}$.
Validation	Demonstrating the successful comparison of the model against data.
Verification	Demonstrating the ability of a model to solve the governing equation for specific conditions.

Common terms which are given different nomenclature outside this thesis

Here	Elsewhere	Comment
κ	α , ω	MIM exchange term is often written as α or ω but it is written here as κ to distinguish it from hydraulic dispersivity
C_m [ML^{-3}]	C_f [ML^{-3}]	Fracture subscript not preferred here because the mobile pathway may or may not look like a 'fracture' in a rock
C_{im} [ML^{-3}]	C_m [ML^{-3}]	Potentially confusable with 'm' for matrix.
D	D_L	The systems examined here are handled in a 1D sense, so no subscript is used to distinguish between longitudinal and transverse dispersion, with D assumed to be the same as D_L
q	J	v in Freeze & Cherry (1975)
$V = q/\theta$		$\overline{V}$ in Freeze & Cherry (1975)
s	p	Laplace variable
σ	$\sigma = \frac{\theta_{im}}{\theta_m}$	In the US groundwater and soils literature $\beta = \theta_m / \theta$, also in the Geotechnical literature, voids ratio $e = V_v / V_s$, where V_v is volume of voids and V_s is volume of solids
Sorption		Can be used to mean 'take-up' of liquid or solute into porous blocks by capillary action or diffusion, rather than infer any surface interaction

Symbols

Symbol/operator	Definition
$\langle x \rangle$	Expectation
f'(x)	f'(x)=∂f / ∂x
x_y	x_y may be written x_y to denote subscript
$\overline{x}$	Laplace transformed variable
$\hat{x}$	Mean
$\underline{x}$ or $\underset{\sim}{x}$	Vector
$\nabla^2 x$	Laplacian

1 Introduction

1.1 Nature of the problem

There is a need to understand transport in porous media (both natural and man-made), in order to predict and then manage the fate of contaminants. By their very nature porous media are difficult to visualise or measure and are often poorly characterized. The most ubiquitous form of interrogation of transport comes from solute concentrations in tracer breakthrough curves (BTCs) and less commonly from flushing data. Characterization beyond concentration-based data can be scarce, expensive and may be difficult to translate into meaningful parameters of a model. This deeper characterization is entirely desirable to understanding transport. However, the assumption in this thesis is of systems in which this further data are either missing or unreliable. For example, there are many streams of waste materials, such as Municipal Solid Waste (MSW), for which this assumption is unfortunately true.

Despite measurement paucity there is a wealth of mass transport models available for porous media covering a wide spectrum of complexity and incorporating a range of different processes. A modeller faces a difficult challenge of balancing a practical and aesthetic desire to keep models as parsimonious as possible, with the risk that an overly simple model may represent the system inappropriately and/or give an unrealistic prediction. In heterogeneous media reliable prediction is not a matter of 'bolting together' 'well-established' processes, because none of the simple models necessarily describe the system well. Instead, there is a delicate path to be trodden in finding appropriate simplicity from complexity (Wainwright and Mullighan, 2004).

In order to move towards better prediction in such systems an analysis of what temporal concentration data will reveal is needed. The result depends on the models used. So, what is required first is an appraisal and organization of the broad range of models available, followed by a selection of a manageable sub-set of these for regular use in prediction.

Amongst the existing models that can be considered for the sub-set, simple physically-based models stand out as being both very common and having a good theoretical basis for their application. There is a continued case for using such models, which includes the substantial associated benefit of comparison with a large number of existing studies that have used the same models.

A structured approach to the sub-set of models allows consolidation of the considerable (but dispersed) attention to these models in the literature and provides a level of organization that allows a sharpened deductive methodology to be applied to model diagnosis.

Consider for example an analysis of exceptionally good data for Municipal Solid Waste Incinerator (MSWI) bottom ash by Beaven, Woodman and Barker (2005). Near identical and 'good' r^2 values (>0.85) were achieved by fitting a 'pool' of transport models to chloride flushing data at two different scales. Noise and a lack of any strong characteristic features in the data meant no single model could be identified as best, yet extrapolation to future time showed a wide divergence of prediction (Figure 1.1 and Figure 1.2). Transport processes in wastes are poorly understood, so there is fertile ground for work on diagnostics of processes in this area.

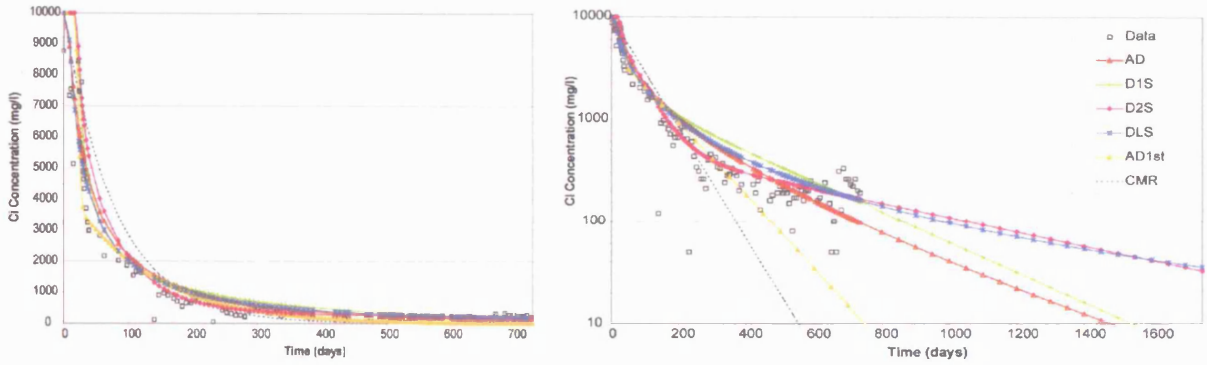


Figure 1.1: Comparison from Beaven, Woodman and Barker (2005) of different models against Chloride flushing data from MSWI ash in a lysimeter from Stegemann et al. (1995): a) Fit to data; b) Extrapolation on log-lin axes. Models in legend are; AD=Advection Dispersion, D1S=Dual Porosity (DP), with spherical blocks, D2S=DP (with two size categories of spherical blocks), DLS=DP (with a lognormal distribution of spherical blocks), AD1st (AD with a mobile-immobile first-order exchange term), CMR=continuously mixed reactor, i.e. exponential model.

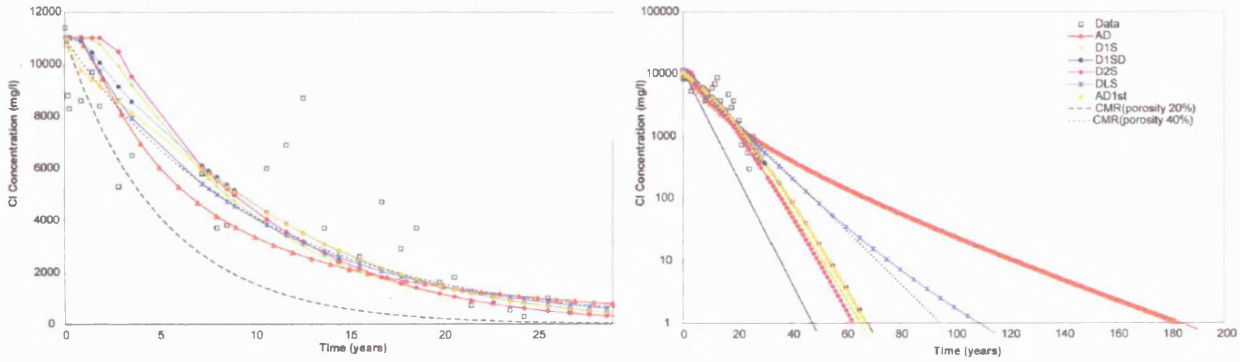


Figure 1.2: Comparison from Beaven, Woodman and Barker (2005) of different models against Chloride flushing data from MSWI ash in a pilot landfill from Hjelmar and Hansen (2004), a) Fit to data; b) Extrapolation on log-lin axes. Legend symbols as per Figure 1.1.

Such problems occur even in relatively well-characterised systems. For example, Moreno and Neretnieks (1985) demonstrated that although good agreement to data was obtained for two different mechanical dispersion models (stochastic-advective and classical advection dispersion) in conjunction with the dual-porosity process, prediction was shown to be strongly dependent on which model was selected. This shows that in order to better constrain prediction, the

conceptual uncertainty (i.e. choice of model) can play an important role in addition to *parameter* uncertainty.

What follows from a clearer understanding of relative model performance is an ability to improve experiments, so that the next set of results will further narrow the parameter and conceptual uncertainty. However, where additional experiments are not possible, or the need to predict is immediate, there is a need to look at the question of how best to predict based on the limited data that there are. Furthermore, the spectrum of models that is currently used still needs to be supplemented by fresh models inspired by data or by assumptions that are as reasonable as those that formed the models that already exist. This stretches the conceptual range, potentially giving new insights and potentially affecting the 'worst-case' envelope of prediction, which may be of great importance.

It is therefore necessary to both consolidate well-used models and methods coherently and to maintain a fertile attitude to the creation of new and challenging models where possible. This thesis is predominantly about the former activity, both in a general hypothetical way and in relation to the specific example of transport in wastes. Indeed, the desire to make statements about processes and parameters in studying wastes *stimulated* the more general discussion.

1.2 Aims and Objectives

1.2.1 Aims of this thesis

The **primary** aim is the improved prediction of contaminant flushing from highly heterogeneous materials. The **secondary** aim is to specifically improve prediction of contaminant flushing from wastes.

1.2.2 Objectives

A series of objectives are required to move towards the aims:

Models

1. To appraise, assess and organise appropriate flow and transport models from a pool of possibilities;
2. To develop a comparative framework of a sub-set of appropriate models for diagnosis of simple processes;

Data

3. To assess existing datasets for transport in wastes at a variety of scales and to discuss the limits of information content;

Future approaches

4. To consider the design of new experiments for better process diagnosis and better parameter recovery;
5. To begin development of fruitful potential entrants to the 'pool' of useful model tools by a model study of one possible stochastic representation (channel network model).

Figure 1.3 puts these aims and objectives in the context of a research programme.

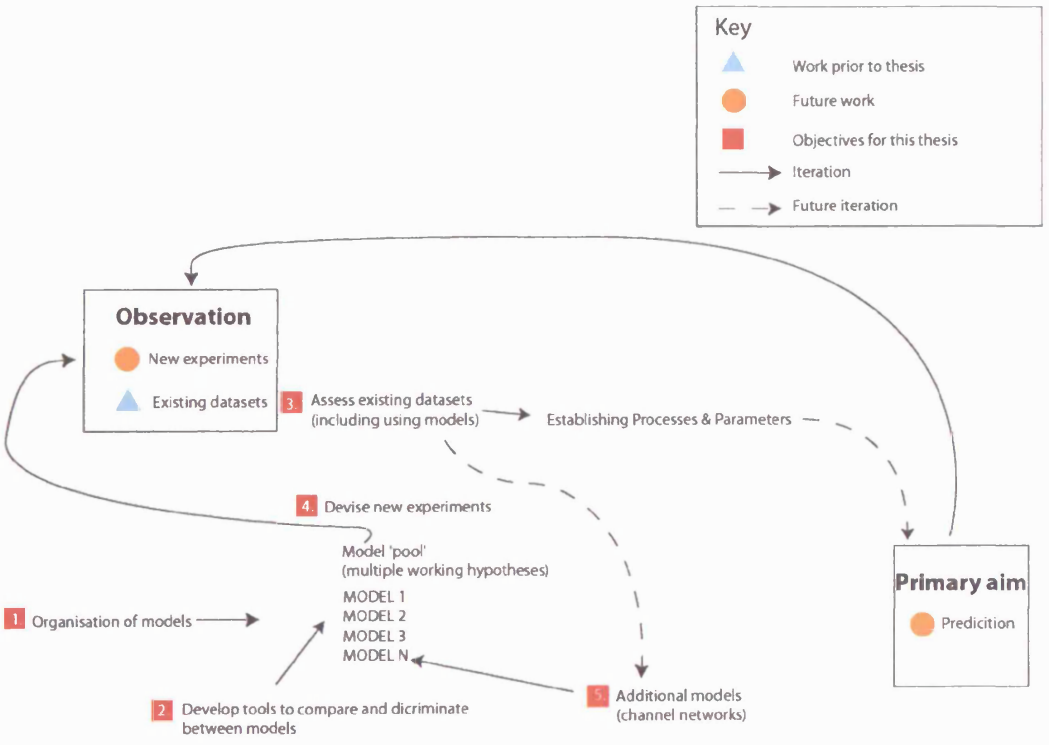


Figure 1.3: Thesis aims and objectives in the context of a research programme.

In generic terms, the objectives of this thesis are all concerned with how to go about modelling and analysing flow and transport data (typically available as tracer breakthrough or flushing concentrations) for highly heterogeneous systems for which little detail is known of the structure or flow geometry within. In more specific terms the thesis can be considered an early phase in iteratively working towards reliable prediction of transport in wastes.

Of primary interest for wastes is the ability to predict flushing. Because tracer tests are a good way of investigating transport properties BTC data is used here in addition to flushing data. In a generic discussion, both forms of data are of interest as they simulate

different problems – firstly of the fate of a solute that is input at a point in the system and secondly of elution of a solute initially distributed within the medium.

1.3 Outline of Thesis

Following this introduction and discussion in **Chapter 2** of the philosophy of the approach adopted, the thesis is divided into three parts in logical succession.

Part I is a generic (theoretical) treatment of the problem of diagnosis of BTCs and flushing data. **Chapter 3** is a literature review for mass transport in heterogeneous systems. A structure is given to the literature and is retained throughout the remainder of the thesis. **Chapter 4** considers methods available for analysing the models and **Chapter 5** develops a flexible framework for generating a wide range of models from a single set of equations. This range of models contributes to the sub-set of models that are going to be more deeply analysed. These models and the methods in **Chapter 5** are simple and well-used, but there are several aspects which are less well-known. Each of the sub-set of models for deeper analysis is examined, in turn, in **Chapter 5** by each of the chosen methods. Single processes as well as process pairs are considered. Scattered literature results are assimilated where available and where no contribution is found new results are presented (or the gap is identified). All this analysis is brought together in **Chapter 6** which compares the different models against each of the measurement methods. **Chapter 7** provides a more concise analysis of the next level in the structure of models (at a higher level of heterogeneity). **Chapter 8** summarises the findings of **Part I**.

Part II concerns transport through wastes. **Chapter 9** is a literature review appraising datasets, tracer experiments and modelling development. **Chapter 10** is an analysis of a lysimeter dataset and **Chapter 11** analyses a laboratory dataset. The lessons learnt from this analysis are applied in **Chapter 12** which presents suggested improvements to waste experiments. Credit and responsibility for the pre-existing data belongs to the researchers at the University of Southampton who collected it; Richard Beaven, Andrew Hudson (Pitsea) and Layi Oni (laboratory).

Part III (which consists of **Chapter 13)** analyses the behaviour of an entirely new stochastic channel network model. The discussion is in terms of connectivity and flow and is performed in the absence of data. The patterns of behaviours that emerge are related to changes to a simple stochastic generator.

Chapter 14 combines the key results of the thesis in a summary of work and conclusion.

Figure 1.4 is a schematic of the thesis structure. The schematic shows how the organisation of the thesis splits into a theoretical treatment and an application to wastes. Each treatment is dealt with by a literature review followed by increasingly heterogeneous representations (in the order lumped models, locally stochastic, spatially stochastic). There is no locally stochastic analysis of wastes, because the existing data do not support use of such models. There is a synthesis of both the theoretical and applied sections in terms of what has been learnt about transport processes and what could be done in future to improve diagnosis. This schematic is used as a 'road-map' and is repeated at the start of each chapter.

1.4 Summary

Table 1.1 summarises the location objectives of the thesis.

Table 1.1: Where objectives are addressed in the thesis.

Objective	**Aims** **1: Generic** **2: Wastes**	**Part and Chapter(s)**
1: To appraise, assess and organise appropriate models from a pool of possible forms.	1 & 2	Part I Chapter 3
2: To develop a comparative framework of models for better diagnosis of simple processes.	1 & 2	Part I Chapter 5
3: To assess existing datasets for transport in wastes at a variety of scales and to discuss the limits of information content.	2	Part II Chapters 9, 10 & 11
4: To consider the design of new experiments for better process diagnosis and better parameter recovery.	1 & 2	Parts I & II Chapter 8 & Chapter 12
5: To begin development of fruitful potential entrants to the 'pool' of useful model tools by a theoretical discussion of one possible stochastic representation.	1 & 2	Part III Chapter 13

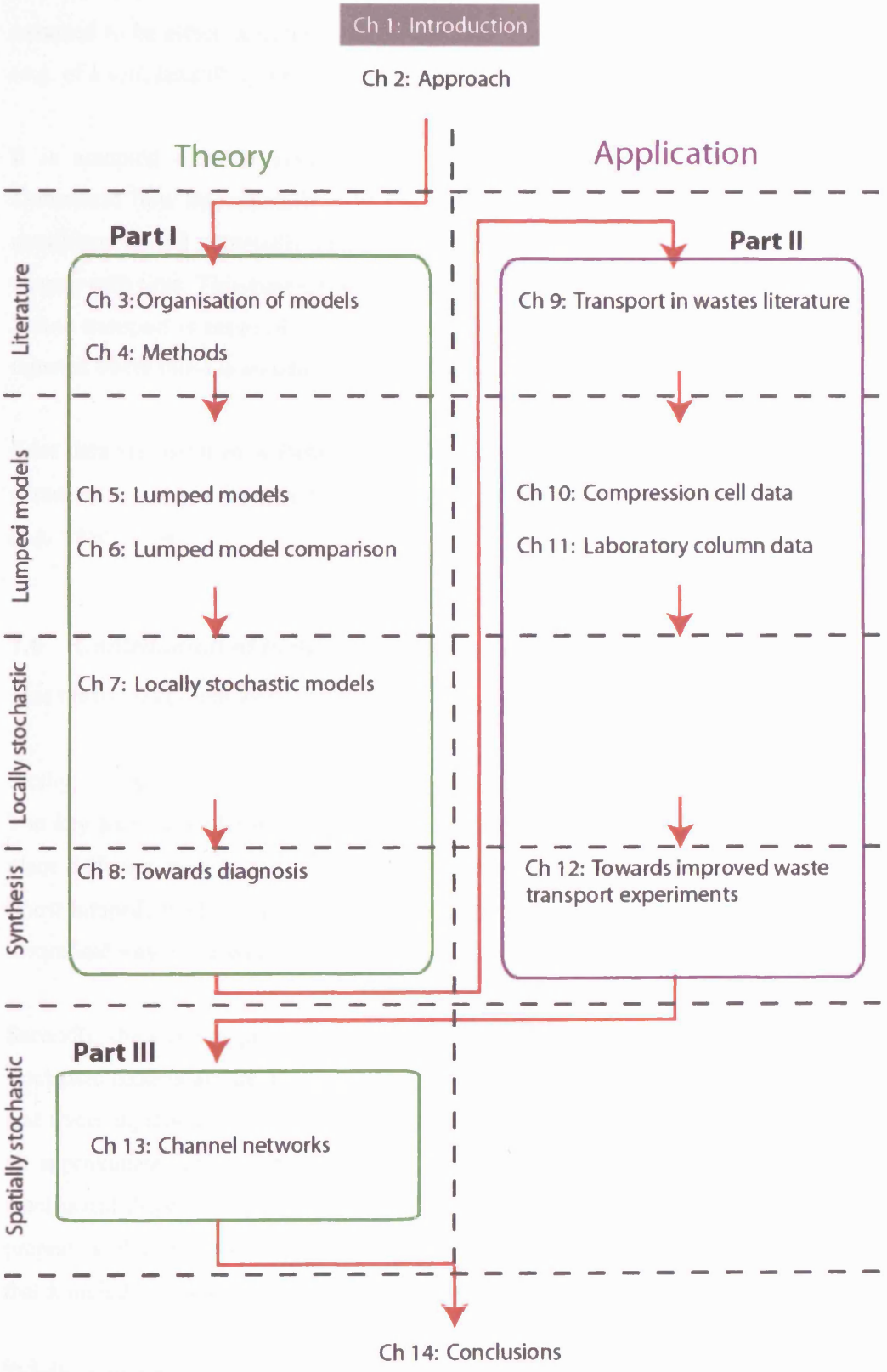


Figure 1.4: Thesis summary.

1.5 Assumptions

The key assumption that runs through the thesis is that the systems being examined are poorly characterised. This means that little is known about the internal structure of the solid medium

and less still about the detailed internal flow and transport pattern. The primary data source is assumed to be either breakthrough data or concentrations of flushed contaminants at the outlet (e.g. of a soil, landfill, lysimeter or laboratory column).

It is assumed that the system can be reasonably represented as one dimensional (1D). Downward flow through soils or wastes, column experiments and certain special groundwater conditions are all potentially amenable to this assumption. The structure is not assumed to change with time. This assumption may, however, be breached by decaying and gaseous wastes. Solute transport is assumed to be conservative (i.e. it is non-reactive and non-decaying) and an injected tracer pulse is assumed to give a unimodal arrival.

Prior data are assumed in Parts I and II (i.e. that there is already some previous knowledge of parameters, perhaps from earlier experiments). Further assumptions are made in specific cases (e.g. steady flow).

1.6 Contribution of thesis to understanding heterogeneous porous media

This thesis makes four key contributions.

Firstly, a new structure is developed for organising simple physically-based process models. The key axis of model discrimination is the level of heterogeneity. This structure is used to place different models and then to find diagnostic characteristics for the least heterogeneous (most lumped) models and for 'locally stochastic' models. This analysis is performed both in a theoretical way using synthetic data and is also applied to new data for transport in wastes.

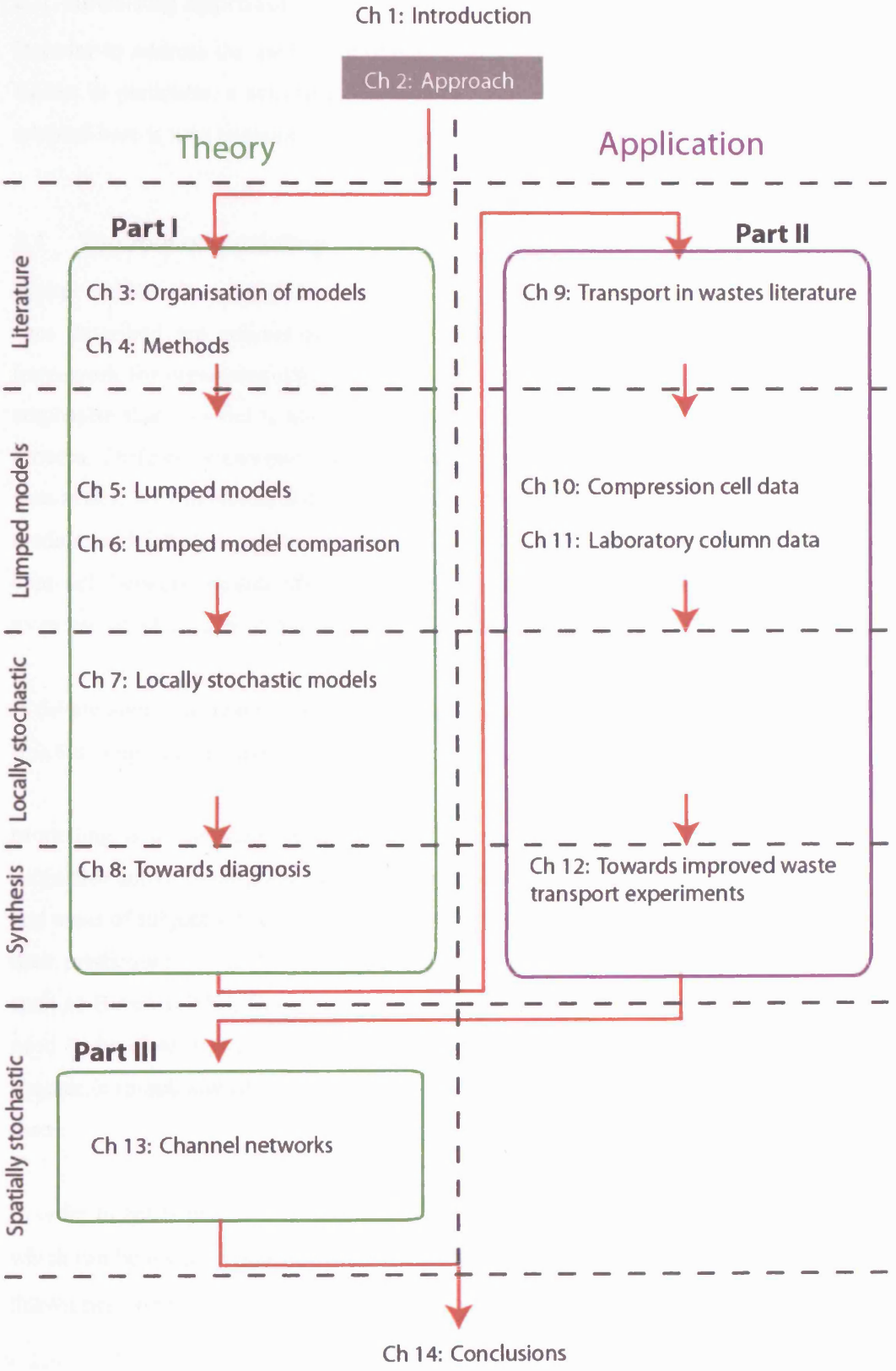
Secondly, there is a range of modelling outcomes. A suite of simple process models and locally stochastic models are developed and their properties analysed. Two piecewise constant flow and tracer injection semi-analytical transport codes are developed - one for dual-porosity with an approximate representation of mechanical dispersion and one fully analytical solution for mechanical dispersion. At the most heterogeneous level in the structure, the network and flow properties of a hitherto unanalysed channel network model are analysed. The model is novel in that it includes spatial correlation in the generation of the channels.

Thirdly, new tracer data in wastes are analysed for the first time. The lessons learned from this data partly motivate the need for the work in the thesis on better diagnostic approaches and enhanced experiments.

Fourthly, there are advances in how to go about optimising tracer experiments. Practical methods are developed to enhance parameter recovery as well as process diagnosis and a range of specific improvements to column experiments for transport in wastes are suggested, along with more general suggestions for advancing the study of transport in wastes.

These contributions are spread through the thesis and are set amongst the considerable existing literature. The extensive attention this area has received reflects both the importance and the number of unanswered questions. The structured approach used here has identified gaps, several of which have been filled. The structure contribution forms the backbone of the thesis and is first developed in **Chapter 3**. In **Chapter 5** the Monte Carlo method is applied to calibration of several simple process models at the lowest heterogeneity level in the structure. Moments and late time gradients are evaluated in a generalised way, making use of the flexibility given by 'Block Geometry Functions'. The importance of the choice of parameterisation is discussed, yielding the surprising finding that very rapid tests might theoretically be able to adequately characterise dual-porosity systems. A simple analytical and numerical exploration is presented of putative Freundlich sorption at low concentrations. **Chapter 6** provides a direct comparison between the process models which were examined in **Chapter 5** (this comparison both draws in literature contributions and adds fresh findings). **Chapter 7** continues the structure at a more heterogeneous level providing a unified and general mathematical approach that combines and reparameterises several rather disparate components that are already in the literature. The analysis in **Chapter 6** and the further work in **Chapter 7** is summarised in **Chapter 8**, which is designed to provide the basis for a practical diagnostic toolkit for tracer experiments. **Chapter 9** summarises the literature on tracers and modelling of transport in wastes. **Chapters 10** and **11** contain new analysis and modelling of waste data that have come from experiments by the liquid and gas flow processes in waste group at the University of Southampton. **Chapter 12** suggests ways in which these experiments should be improved in the future. Finally, **Chapter 13** presents an entirely new set of modelling results using an existing channel network model.

Since several of the contributions are either built on existing work (such as the data analysis of the waste tracer experiments) or are presented here amongst critical literature review material, the nature of the new contribution of this thesis is detailed at the start of each chapter.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

2 Modelling approach

In order to address the aims of improving prediction of flushing generally and of flushing in wastes in particular, a scientific approach is needed. The selection of the specific approach adopted here is now examined in this chapter.

2.1 *The role of modelling*

Barker and Kinniburgh (1994) described a wide range of possible uses for models. Amongst the uses described are conceptualisation aid, thinking tool (hypothesis testing and sensitivity), framework for organising data, framework for collecting data and parameter estimation. They emphasise that a model is not a substitute for data collection or the 'human' decision-making process. Different views persist over the use and meaning of common physical models and their parameters, how to handle heterogeneity in modelling and measurement, how complex to make models, and whether and how to incorporate stochastic processes in models. Difference in approach between studies may relate to different aims or objectives but may also relate to a more profound difference in modelling philosophy or methodology.

A debate about the nature of modelling exists in recent hydrological and groundwater literature. It is thus important to make clear what approach is being used here.

Modelling is at the heart of the scientific process, yet can be at times a highly 'artful' and subjective activity masquerading as objective truth. There is a clear need to make assumptions and areas of subjectivity explicit, and to be explicit about the true limitations of the models and their predictions. This does not always come through strongly in certain studies and authors such as Beven (1996), Gupta et al. (1998) and Konikow and Bredehoeft (1992) highlight the need to be clear about how far a particular model can reasonably go. An intention of this chapter is to acknowledge these issues and demonstrate how this thesis attempts to deal with them.

In order to better predict transport in complex systems, models are considered to be *hypotheses* which can be tested in order to refine the judgement of what is the best model (no distinction is drawn here between a model and a hypothesis, as per Hilborn and Mangel (1997)).

2.2 *Choice of process models*

Physically-based process models are chosen for use in this thesis due to three primary advantages of them over empirical models. Firstly, process models are representations of hypothesised physical behaviour. Empirical models, on the other hand, are normally little more

than functions that appear to give fairly similar results to the measured data.[1] Secondly, a process model is likely to have parameters which can potentially be compared to reality (or independently measured). Thirdly, process models may be able to make other predictions. For example, they may be able to predict the spatial distribution of concentration based on calibration of concentration against time at one point or potentially extrapolate to make sensible predictions when the process conditions are changed from what they were during the experiment.

The predictive benefit of the third point may only be taken so far. Models may only be useful for a window of states but begin to become inaccurate outside that. There is therefore a danger of using, for example, a model calibrated under a relatively slow flux rate, under a much faster one.

If outputs from all combinations of variables could be measured, then there would be little need for a model in terms of predicting the expected output at any time (because it would be possible to just interpolate from a response surface). A model may still be useful in such an instance as an efficient 'curve' fit and also for gaining an understanding.

There is considerable material in the literature analysing why or when complex systems 'collapse' to certain common 'end-member' process models. [2] One example is Matheron and De Marsily (1980) showing that the Advection Dispersion (AD) model can be derived for flow that is not exactly parallel to stratification.

Transfer functions are commonly used for describing process models in 1D transport (for example Becker and Charbeneau, 2000; Jury and Roth, 1990; Sardin et al., 1991; and White et al., 1986). A transfer function is a Laplace-Transformed function which describes the output divided by the input to a system.

[1] Empirical models can still be very useful and may have links to underlying processes. A lognormal distribution of travel times for instance might just be an empirical guess, but there may be physical reasons why a lognormal distribution may arise in the presence of multiplicative random processes.

[2] 'End-member' is defined here as a process model that is a limiting (and simplified) case of a higher-order (more complex) model. This could be due to the asymptotic approximation of a stochastic model by a partial differential equation (pde), by a parameter tending to zero or infinity or by one process being 'crossed out'. The specific interpretation therefore depends upon context. Although all 'end-members' are extreme in some sense in terms of assumptions about the processes, this does not automatically imply that the predictions are extreme in any sense. Although potentially derived from a more complex model, an 'end-member' is not necessarily the simplest model (it may be that further simplifications involving other parameters and processes remain possible).

2.3 *Modelling methodology*

A common view of modelling is as a chain of activities (Anderson and Woessner, 1992; Barker and Kinniburgh, 1994; Wainwright and Mullighan, 2004): problem → conceptual model → mathematical model → coding → verification → calibration → validation → sensitivity/uncertainty analysis → output (prediction).

Iterative loops through parts or all of this chain are inherent to the process. The 'science' part of the process – a loop back from calibration to conceptual model – is the primary interest for this thesis. The details of models, coding and verification that occur for this thesis are presented in Appendix 1 in order to give focus to the conceptual issues in the main text.

The definition of each of these activities may vary from study to study, so the meaning interpreted here is now clarified.

Verification is defined here as the testing of numerical models against known solutions (most likely analytical but possibly by other numerical models). Barker and Kinniburgh (1994) suggest that verification tests are best done over a wide range of parameter values, so as to be most demanding on the code. There are difficulties with the concept of verification (Konikow and Bredehoeft, 1992; Oreskes et al., 1994), because the point of a numerical model is to stray beyond the simple boundary conditions under which it can be 'verified' against an analytical solution. Verification is therefore more of an eliminatory hurdle, for which the model is known to be wrong if it fails, but if it passes it might still be wrong under conditions away from test.

Calibration is defined here as the determination of model parameters by adjustment in order to provide a 'reasonable' fit to data. Barker and Kinniburgh (1994) use a wider definition that allows the boundary conditions and driving stresses to be varied. There are a few decisions that must be made in order to do this. Measures of what constitutes a reasonable fit must be decided on. The best fit can be sought with an automatic fitting routine. Such tools need to be used with the knowledge that there will be a surface of goodness-of-fit with not only a set of fit points around the stated best-fit but with the possibility of other local best-fits and regions or directions of similar fit.

Validation is defined here as the demonstration of a successful comparison of the model with data that were not involved in the calibration. The comparison of the different predictions can give a measure of how uncertainty increases at greater 'distance' from the point of calibration. Konikow and Bredehoeft (1992) pointed out that the oil industry literature talks of 'history matching', rather than model calibration or validation. Konikow and Bredehoeft (1992) further dismissed the notion that there ever can be validation of a groundwater model, showing that this

would require a Positivist rather than Popperian stance in the first place. However, validation is variously defined and Maloszewski and Zuber (1992, 1993), defined it as a qualitative process, based on judgment. De Marsily et al. (1992) countered these views by considering that successful validation shows that the model is not invalidated by the data thus far. Perhaps the best way to think of this process is of confirmation (but not validation) (Beven, 2005; Oreskes et al., 1994). Bridging what becomes a rather semantic gap, Wainwright and Mullighan (2004) proposed the term 'conditional validation'.

Sensitivity analysis, defined in the way described by Anderson and Woessner (1992) is best suited to determining data collection strategies before any data has been collected. When data and prior estimates of parameters and their uncertainties already exist through calibration of previous data sets, '**uncertainty**' analysis would be more appropriate. 'Uncertainty analysis' is used here to mean assessing the effect on the output due to the likely errors within the parameters. Varying each parameter by either a relative or absolute amount about the point of best fit in a sensitivity analysis supposedly uncovers which parameter is 'most important' in affecting the output. However, the most sensitive parameter is less important if in practice it is already very tightly constrained. There is a problem that possible correlation between parameters is not taken into account this way and a false impression of sensitivity is given. More sophisticated interpretations of sensitivity analysis exist but are beyond this discussion (Saltelli et al., 2000). Since some prior data are assumed in Parts I and II, sensitivity analysis is considered inappropriate.

2.4 *Hypothesis testing*

Calibration and validation inherently test models (hypotheses) against data, but how exactly should this be done?

Young et al. (2004) argued that choosing a set of process models (to eliminate deductively) in advance of the data is arbitrary. Instead 'inductive' methods such as the Data Based Mechanistic (DBM) modelling of Young et al. (2004) could be employed. Young et al. (2004) illustrated DBM by induction of a two-rate first-order ('MIM') transfer model for modeling tracer dispersion in a river. However, in defence of selecting simple process models *a priori*:

- Such models have multiple explanations for how they end up as end-members, so their forms are not whimsical constructs.
- Models that are highly-used may be so because they reasonably predict across broad ranges of systems (but also may be used through 'fashion').
- Selecting multiple models carefully distributed over a range of processes and heterogeneity levels avoids the risk of adopting the 'pet' model.

This thesis therefore remains deductive although inductive methods may have important application in the future, particularly for prediction. The set of models to be tested is here termed a 'pool' of models (i.e. the same as Multiple Working Hypotheses, as first described by Chamberlain, 1897).[3]

Calibrating all the models against an objective function and selecting the model of best-fit would be one way of selecting the 'best' model, although it is possible that several models achieve similar best scores.

Instead it would be better to choose a rejection criterion (or criteria). If a model only achieves a 'poor' fit to data then the model itself could be rejected and a different or modified model suggested. This process would be Popperian in nature (Popper, 1979). The question is then how to set a rejection criterion. For BTC data there is no standard criterion and the choice is kept subjective (conventional hypothesis testing is rarely, if ever, used for transport model selection). One reason for this is that there is a wide range of requirements for models to address (for example, predicting the maximum concentration, or the total mass that will have arrived at a given location).

When analysing goodness-of-fit, the presence of noise in the data must to be taken into account. For example, if using r^2 as a criterion, noise will set an upper bound on the best possible r^2 that could be achieved even by a 'perfect' model so there would be no point in setting the threshold above this level. Freer et al. (1996) varied the choice of rejection of 'non-behavioural' simulations trying $r^2>0.3$ and $r^2>0.5$ and Beven and Freer (2001) used $E_N>0.6$, where E_N is the Nash Sutcliffe efficiency. Simulations which exceed this threshold are considered to be 'behavioural'. Beven (2002) called such low rejection criteria a 'relaxed form of Popperian falsification'.

It may become apparent that possible fits are distributed in a continuum which is to be cut by this arbitrary threshold. Setting a high threshold might result in the rejection of all models (as they are all in some senses approximate). Setting the threshold low can result in acceptance of all models and failure to discriminate. It is appreciation of this sort of behaviour that led to the sensitivity method of Spear and Hornberger (1980). The cumulative distributions for each parameter were compared for sets of 'behavioural' models (ranked by the fits) giving an indication of the sensitivity of the parameter.

[3] Chamberlain (1897) argued that having only a single 'pet' ruling hypothesis is to be avoided.

Different goodness-of-fit measures will need different rejection criteria and may well result in different outcomes in terms of model rejection. Rejection can also take place through other (non-BTC) knowledge of the system.

An inability to match certain effects may not imply an inability to accurately predict the phenomenon of interest. For example, it is not clear whether the inability of a simple steady-flow dual-porosity code to cope with short-term saturation transients impairs the ability of such a code to predict long-term flushing. Nor is it clear whether failure to model a concentration change during a pause in flow means that the advection-dispersion model should be completely rejected.

Failure for parameters to relate correctly to what may be measured may be an indication that there is some form of model structural error, or that the parameter measurement is not commensurate with the averaging scale of the model (an example of an unrealistic parameter would be porosity of 80% in a sandstone rock). Either way, a model with a 'parameter out of range' may still reasonably mimic calibration data and possibly provide useful prediction. It may be acceptable in some circumstances to employ 'apparent' parameters that lump various processes together, that are difficult to distinguish (Maloszewski et al., 1995). Given that models are known to be approximate to some degree, a distinction is put succinctly by Beven (2002) between predictive power and explanatory depth; since predictive power is the aim here, it may be acceptable to sacrifice explanatory depth.

Rejection due to unrealistic parameters is potentially difficult. Physical measurement may be extremely difficult, particularly due to the presence of heterogeneity. Some parameters (such as dispersivity) may be unmeasurable by any other means than by calibration. An example of the lack of robustness of using parameter values to eliminate whole models is the argument made by Mathias (2005), who used limited data on the size of chalk blocks (which may anyway be only partially related to the flow geometry) to contribute to the invalidation of a dual porosity (DP) model for chalk. The analysis relied on an assumed fracture aperture which remained fixed, despite the variability of block size. The total fracture porosity was therefore not being kept constant. Arguably the case by Mathias (2005) should be used to promote a new entrant to a pool of models, but is insufficient without further data to reject one model in order to replace it with another.

Consideration of rejection of models may also need to take into account model complexity, particularly because the more parameters held by a model the better it can be tuned, but at the same time non-uniqueness can set in. There are methods for how to handle whether or not additional parameters added to hierarchical (nested) models should be accepted as an improved

model, which essentially penalises the goodness-of-fit for introduction of new parameters (Efron and Tibshirani, 1993; Linhart and Zucchini, 1986). However, the models used here are not all nested in this way. The MultiModel Analysis part of the USGS 'JUPITER API' includes a number of methods.[4]

Like all models, models of heterogeneous transport are only approximate. The selection of models therefore involves questions over the degree of approximation that is adequate for use in prediction and how much less realistic a model is to a 'competitor' before it is completely rejected. In practical terms, efforts towards prediction can and must be made, even before the process of rejection has converged on an "explanation and solution" (Schumm, 1991). Such a final goal is not necessarily the same as having found the 'truth', which is possibly unreachable anyway; Beven and Binley (1992) asserted that this is the case for distributed hydrological systems.

A potential mistake is the rejection of a wider hypothesis on the basis of a specific model. For example, Becker and Shapiro (2000) eliminated the possibility that diffusive mass-transfer is responsible for breakthrough tailing on the basis that a simple dual-porosity model fails to adequately model the data. By 'elimination' they find that advective-dispersive effects are all that remain to explain the data, but do this with awareness that a system with multiple scales of diffusion is a highly feasible alternative.

The creation of simple process models can be thought of as mapping a series of discrete representations to a broad range of possible conditions. The idea of eliminating models from a number of apparently clear choices can helpfully be viewed on this basis.

A better simple representation might be missing from the pool of testable models. Alternatively it may be that all the models in the pool share common assumptions that may be highly questionable (for example the level of heterogeneity assumed).

2.5 *Lack of discrimination between models*

There may be further problems for the concept of hypothesis testing of BTC data. With ever increasing computing power, extensive exploration of error space is possible (Carrera et al., 2005). This exposes potential weaknesses that can befall calibration. The problem is that, particularly for a model with several parameters, there may be a wide distribution of parameter combinations that give equally 'good' fits. Beven and Binley (1992) explored parameter space by sampling from a Monte-Carlo set of realisations based on initially uniform ('prior')

[4] For example, the Akaike Information Criterion (Akaike, 1973), the Bayesian Information criterion (Schwarz, 1978) and Hannan and Quinn's criterion (Hannan and Quinn, 1979).

probability density functions (pdfs). Beven and Binley (1992) called this 'equifinality', and it is illustrated by use of 'dotty' plots showing the goodness-of-fit against parameter value for any particular parameter.[5] Such plots can sometimes exhibit no structure at all, casting great doubt on the concept of a single parameter optimum. Meaningful calibration would be difficult in such an instance.

Beven and Binley (1992) also discussed 'non-uniqueness' which they defined as the existence of multiple minima in parameter space. If gradient methods are used, only one minimum will be found unless the fit is repeated many times from different starting points. If many minima are found, careful thought is required over how they can be used for prediction. They distinguished non-uniqueness from 'non-identifiability', which is defined as a poorly defined parameter optimum.[6] This can occur due to plateaus or lines ('valleys') of indistinguishably good fits which make it very difficult for gradient-based automatic calibration algorithms to converge on a single point.

The findings of Beven (2002) are based originally on work on highly parameterised distributed hydrological models, but have subsequently found more generic application in environmental modelling. Abbaspour et al. (1999) observed multiple minima and valleys in a response surface generated by fitting the MACRO model (Javis, 1991) to landfill flow data. MACRO requires 38 parameters. It is an interesting question as to what extent non-identifiability and non-uniqueness problems pervade simpler porous media models than this.

In the instance of a column test, the system potentially has well-constrained boundary and initial conditions. However, column tests in some media, for example wastes and clays, are difficult to control and therefore there is a higher level uncertainty in the process conditions. Transport in a highly heterogeneous material may be governed by non-linear underlying processes and is likely to have an unknown flow geometry.

The 'equifinality' problem can be shown for simple models applied to certain BTC data. Maloszewski and Zuber (1993) observed the problem for very simple models. Becker and Charbeneau (2000) showed that 1D or radial transport representations gave equally adequate fits to a weak-dipole tracer test. Jury and Roth (1990) demonstrated how three simple models can give very similar BTCs. Chapter 10 demonstrates that equally good fits to data can be achieved

[5] Equifinality can be observed for sets of parameters for a given model but if several models are being tested it can apply to complete model-parameter sets. This cannot not be plotted in a conventional dotty plot, however.

[6] The terminology used varies between authors (for example Kleisson et al. (1990) discussed the related issue in terms of a lack of identifiability between models). The definitions given by Beven (2002) are adopted here.

with different assumptions of the shape of the dual-porosity block geometry. Wright (2002) demonstrated how, although different block geometries show very similar predictions for a period of time, they diverge significantly for longer simulations (the example used was for several repeated aquifer storage and recovery (ASR) cycles).

These observations appear to be at odds with results from typical calibration methods. Zhang et al. (2006) analysed soil column experiments using the CXTFIT code (van Genuchten, 1999), finding for each column a single optimum set of parameters and (fairly narrow) 95% CIs for each. Zhang et al. (2006) does not make the case that the problems observed in distributed models are necessarily applicable to column tests (namely non-identifiability and non-uniqueness). Instead they contend that 'conventional' methods such as CXTFIT make a number of incorrect assumptions, including for example that the error variance is constant and that the model is correct. They find the result of this is that conventional methods substantially underestimate parameter uncertainty and therefore predictive uncertainty. They demonstrate this point by applying the 'Generalised Likelihood Uncertainty Estimation' (GLUE) methodology to the same data. The spread of 'behavioural models' (which they define to be all parameter combinations giving $r^2 > 0.9$) is substantially wider than the 95% CIs calculated by CXTFIT, and the range of predictions is similarly wider. This choice of $r^2 > 0.9$ is highly questionable.

Critics of GLUE find, conversely, that GLUE overestimates the uncertainty (Mantovan and Todini, 2006) and they particularly object to the subjective choice of what constitutes 'behavioural'. So, there is currently no consensus as to the best form of prediction method. What is certainly clear is that although the precise method requires further research, attention has been drawn to the central importance of uncertainty analysis in hydrological prediction (Pappenberger and Beven, 2006) and the need to question the many difficulties inherent in conventional approaches.

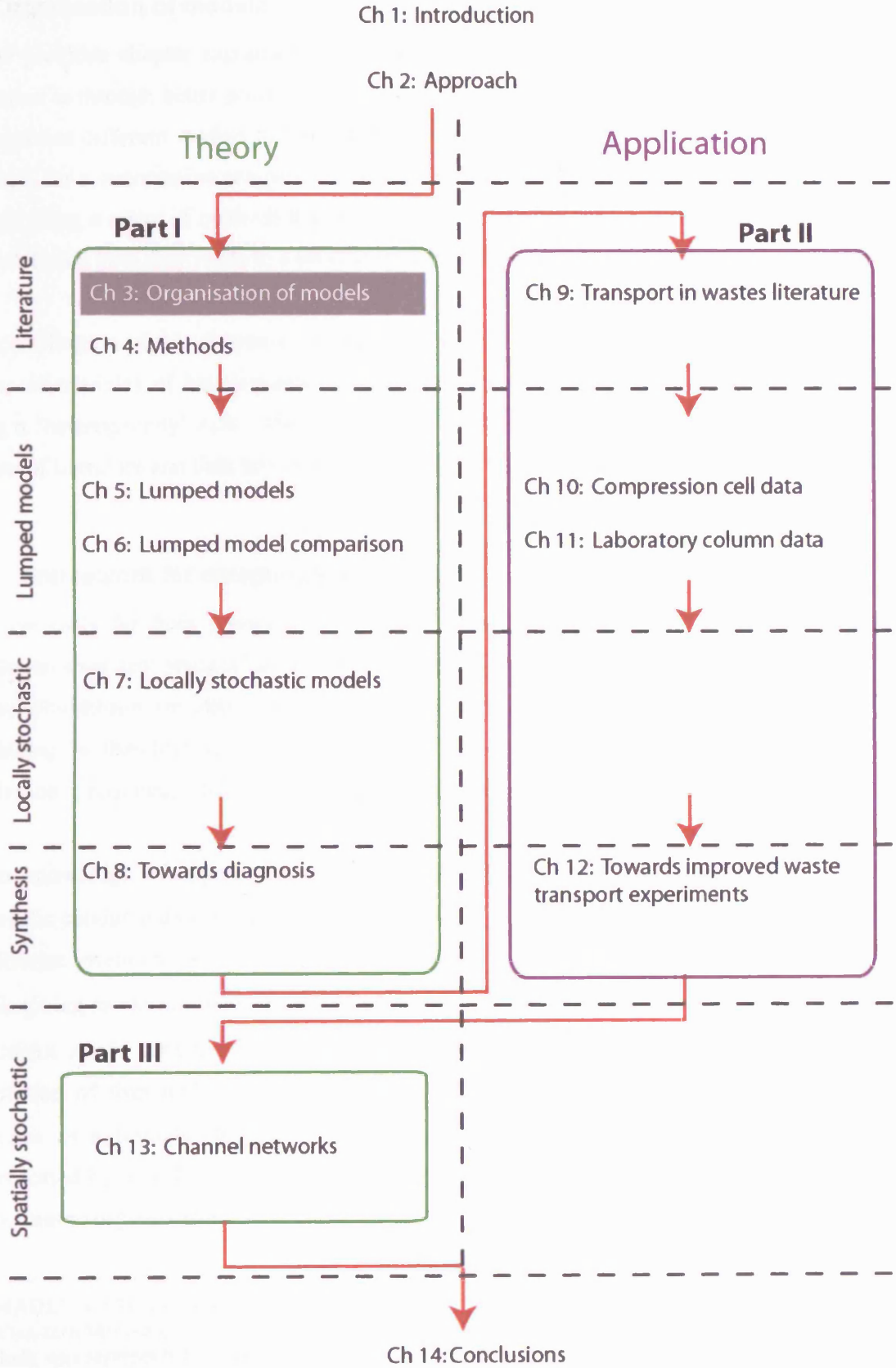
Although establishing the uncertainty of parameters is not necessarily a step on the way to prediction, it potentially is. For example, there are predictive approaches such as the method used by the PEST code (Doherty, 1994), which use the 95% CIs of parameters to establish an extreme range of predictions. Although this thesis stops at the point of addressing parameter and model identification, the findings will potentially have a bearing on prediction in the future. Stopping at the identification point is a very useful place for examining scientific inference about transport dynamics. The 'search for an enhanced approximation of the truth' has the additional benefit that the chances of reliable prediction are likely to be improved if the system is better understood.

Although there is evidence to suggest that the discriminatory power of BTC data is potentially quite low, what is needed is a quantitative exploration of the limits to what *can* be known from a given BTC (see Chapters 5 - 7). Where there is the opportunity to control and repeat (in the laboratory) it is worth investigating deductive methods further, by developing more discriminatory experiments. Of course, a BTC is rarely examined in isolation from wider knowledge about the system. There are many further questions raised about how different forms of evidence and preconceptions can also be incorporated into judgements about the plausibility of particular models. However, in this thesis focus is on the information content of a given BTC rather than on how to incorporate this wider knowledge.

2.6 *Method adopted for this work*

Models are used as a pool of multiple working hypotheses. The pool is populated with a range of well-established plausible, physically-based 'end-member' process models. These models are discussed and organised in Chapter 3. It is desirable to reject hypotheses due to failure to represent key behaviours (with plausible parameters). The nature of rejection is shown to be a function of a subjectively chosen measure, but can also include logical tests which don't need to suffer the 'shades of grey' rejection based on goodness-of-fit. With certain datasets the discrimination between models is potentially poor. There is a need to carefully compare the behaviour of the models in the pool to identify clear predictions which can be used to diagnose different processes in a logical way (see Chapter 6). Since investigation is an iterative process there will still be a practical need to predict before much refinement has occurred. In this instance a sensible approach would involve consideration of Bayesian 'risk' models.

So, the process of hypothesis testing using BTC data is not necessarily objective, straightforward or free from potential pit-falls. However it is more rigorous than arbitrarily building a single model.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 9: Transport in wastes literature
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14:Conclusions

3 Organisation of models

As the previous chapter explained, the chosen route to achieving the main aim of enhanced prediction is through better process understanding. This chapter addresses the need to organise and appraise different models in broad terms (Objective 1). This organisation is then used as the basis for a quantitative comparative analysis (Objective 2) of a selected sub-set of common models using a range of methods that are developed in Chapter 4. The selection of the sub-set of models and how they relate to a broader spectrum of models is also discussed in this chapter.

The contribution of this chapter is to suggest a new way to organise mass transport models by means of triangles of key processes (mechanical dispersion, immobile storage and sorption) along a 'heterogeneity' axis. This structure is applied in this chapter in order to organise the review of literature and then serves to organise all of the remaining chapters.

3.1 *Framework for categorisation*

It is currently far from possible to describe the geometry of a porous system at pore-scale resolution over any practical scales (although Martys and Chen (1996) have done this using a Lattice-Botlzmann simulator for sub-centimetre scale sandstone). A challenge for transport modelling is therefore to capture key characteristics of the system at the scale at which prediction is required, whilst averaging behaviours that take place at smaller scales.

Even knowledge of apparent properties of the system can only be partial. For example, hydraulic conductivity (K) might be sampled at a number of scales at a relatively small number of discrete locations. At the MADE site 2483 borehole flow measurements were taken from 58 wells giving an exceptionally high resolution of $K(x,y,z)$ relative to most aquifer scale studies (Rehfeldt et al., 1992).[7] However, for the plume to be measured 'deterministically' (to a resolution of four nodes per correlation length) 400,000 nodes would have been required. Perhaps as a response to conceptual uncertainty, some highly studied systems have been represented by a wide range of different models (for example, transport in the Chalk).[8] It is important to organise this uncertainty.

[7] 'MADE' stands for 'Macrodispersion experiment', which occurred at Columbus Air Force Base in northeastern Mississippi.

[8] Chalk was represented by the ADE in many early studies. Thereafter a range of mass-transfer models has been employed. Examples include: diffusive transfer models (Bibby, 1981; Gaus et al., 2000), MIM transfer models (Watson, 2004); spheres with fracture skin (Moench, 1995); 1D diffusive models (Fretwell, 1999; Wright, 2002); stochastic discrete fracture network models (Moreno et al., 2003). This is partly explainable through spatial differences as well as changes between the different strata, through modelling occurring at differing scales and for different purposes, by refinement in the commonly agreed conceptual model, and development and availability of new models. Arguably, however, despite a relatively high degree of research and characterisation work, considerable conceptual uncertainty remains. The role of (perhaps karstic) channels for example is becoming increasingly apparent (Maurice et al.,

Although the exact geometry is currently unknowable it is possible to perceive a complex 'reality' from which simple conceptual models may arise. A medium is here imagined to have a complex spatial distribution of permeability and sorbtive properties. A simplified 'perceptual' model is shown at the top of Figure 3.1, where this distribution of permeability gives rise to distributions of interacting flow-paths and, as a corollary, distributions of less permeable ('immobile') zones. To some extent this is a false distinction, but it is a very useful one, since different processes may apply; advection for flow paths and diffusion may be important in less mobile regions. Superposed on to this physical picture are chemical interactions, in particular the interaction with surfaces by sorption. The complex reality is simplified to a 'triangle' of three processes.

There are many possible ways of organising models, for example some literature reviews are structured chronologically, which is interesting in terms of the history of the subject but does not necessarily organise the processes coherently. Barker et al. (1994) provided a comparison of codes based on their functionality. Lichtner (2000) presented a categorisation that does not easily accommodate different complexity scales. Barker (1991) set out a timescale-based 'map' primarily aimed at demonstrating the relationship between diffusional times and block sizes (note that the focus here is on solute transport, not on pressure or heat transients, however, with a minor change in parameters the mathematics would be identical). Model space could be mapped out in terms of a distance scale against a time scale (Woodman, 2004 cited in Wright et al., 2005). There are of course several distance and time scales relevant to heterogeneous systems beyond the notion of block sizes and times. There have been various attempts at combining the key timescales into single dimensionless times beyond 'standard' dimensionless variables which can be thought of as time ratios (for example Peclet number).[9]

A different organisation is suggested here. By perceiving reality as a triangle of individual and coupled processes, it is natural to organize simplified conceptual models of this reality in a similar way (Figure 3.1). A further distinction can be made between advective and dispersive processes (despite the fact that mechanical dispersion is really the result of a range of advective paths). This distinction is very useful for the categorisation of models and the examination of analytical solutions which only include single processes in conjunction with advection (for

2004). In this light, it may be that a broad range of model space is appropriate and the medium is not as conveniently amenable to classification statements such as "the Chalk *is* a dual porosity system" or more recently that it "*is* a dual permeability system". This becomes only a semantic point if '*is*' is translated as shorthand for "*is apparently best modelled*". However, where the more literal interpretation does arise it indicates a potentially dangerous over-estimation of conceptual certainty.

[9] For example, Young and Ball (1995) defined $\tau = t_a \sigma N_p / t_{cb} R$, where N_p is number of pore volumes of fluid input and R is the retardation coefficient.

example sorption and advection in the absence of mechanical dispersion). The basic elements of this categorisation are three primary processes that are arranged in a 'triangle' around the central **advection** process - **mechanical dispersion** (which relies on advection), **immobile exchange** and **sorption**.

The next question is how to cope with the impact of complexity and heterogeneity on conceptual models. Although the simplest possible representations are sought here, complexity is a further axis in 'multidimensional model space'. The many other physical, mechanical, biological and chemical processes that might affect some systems are also best imagined as further axes (out of plane), but these are (by assumption) not needed in the analysis here. Heterogeneity is separate from complexity, as it can be represented in very parsimonious ways (for example, in a lumped model dispersivity (α) can be thought of and used as a 'heterogeneity parameter', and a spatially stochastic representation although producing a complex permeability field could have a highly simple generating algorithm with only a few parameters at its core). The following section discusses heterogeneity.

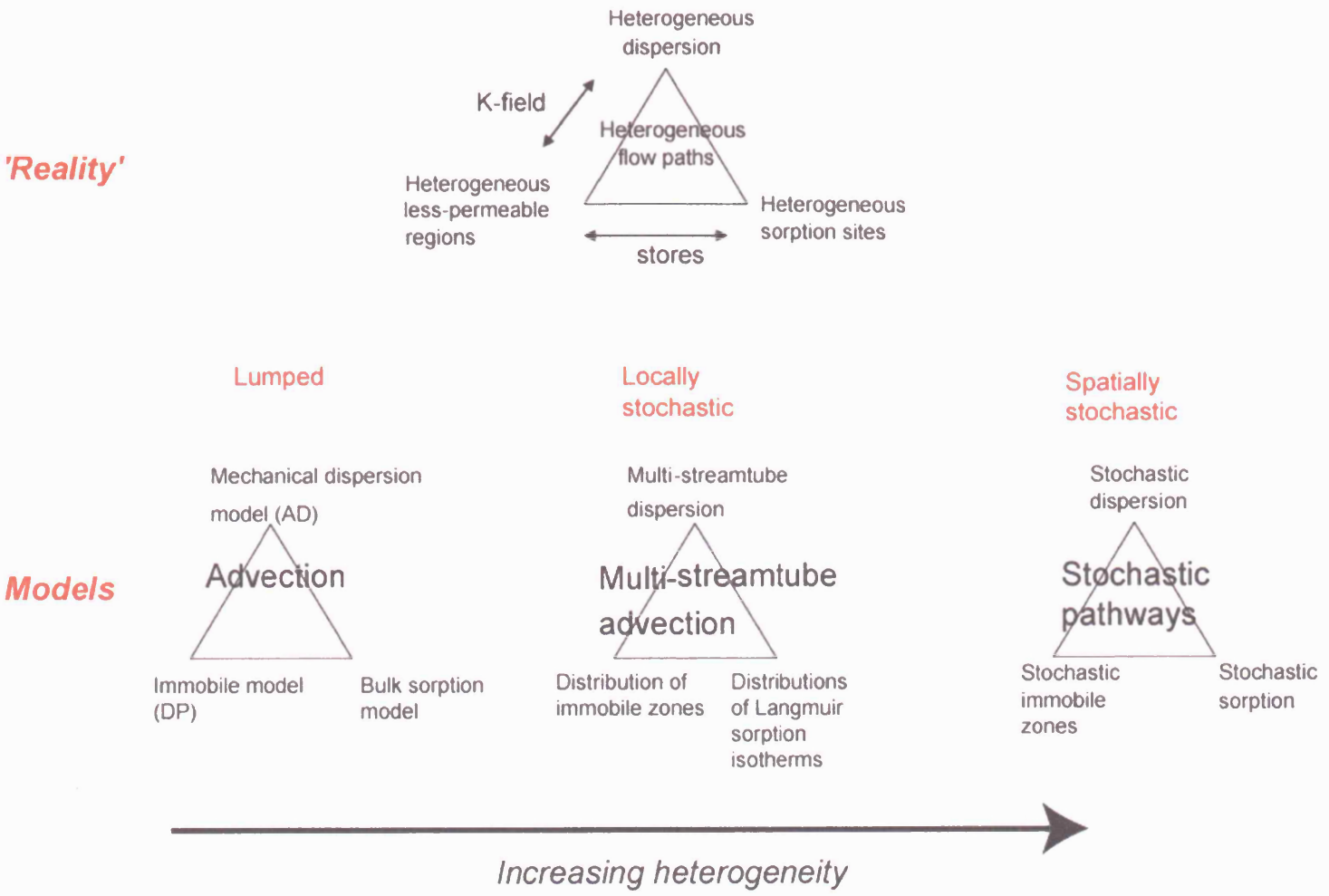


Figure 3.1: Organisation of models.

3.1.1 Heterogeneity axis

Freeze and Cherry (1979) defined a system as *homogeneous* when $K(x,y,z) = C$, where C is a constant and as *heterogeneous* when $K(x,y,z) \neq C$.[10] A key measure of heterogeneity has been the standard deviation (SD) of $\log(K)$, σ_Y (Gelhar, 1986).

Bear (1972) defined the representative elementary volume (REV) as the macroscopic sampling volume at which a measured property does not change. The medium is considered to be a continuum with properties at a point having a value pertaining to the REV centred at that point.

Gelhar (1993) pointed out that this requires a disparity in scale, because the averaging process over the REV are averages of fluid properties which themselves are averages from a molecular to fluid continuum level. Heterogeneity of porous properties at the field scale requires another averaging scale to yield effective macro properties (such as macrodispersivity). In terms of length scales, the dimension of the REV, D, is bracketed by two further length scales, $l \ll D \ll L$, where l is a characteristic microscopic length scale (pore scale) and L is a characteristic macroscopic length scale. Baveye (2004) pointed out that if the upper condition is not reached, a plateau of constant property value may never be reached. Heterogeneities could evolve with scale, as discussed by Cushman (1990). Tsang et al. (1996) noted that if the heterogeneity obeys a multi-scale organisation, and/or channelling occurs, the macrodispersivity may never tend to a limit.

Baveye (2004) found that the REV is an untested hypothesis. He argued that the concept of a fixed REV is a constraint and instead it might be better to think of measured properties as a convolution of microscopic properties with particular weighting functions that are caused by the particular measurement instrument in use. This in principle allows changes in the volume measured, both spatially and temporally, and also even an inclusion of instrument error. The particular example he cited is that of porosity measurements by neutron probes (note that other measurements, such as aquifer pump tests, also sample variable volumes and so are also amenable to this treatment). He went on to calculate macroscopic mass balance equations based on the integrals of the microscopic functions and show that the conventional REV approach is a sub-set of this more general treatment. The fact that Baveye (2004) is still arguing for the

[10] Greenkorn & Kessler (1969) adopted a different definition of homogeneous: that mean K remains constant, i.e. $\overline{K(x,y,z)} = C$. If K is constant everywhere, then the medium is 'uniform'. Both really only have meaning relative to a reference scale and relative to the instrument that measured the property of interest. Freeze (1975) also defined homogeneous to mean a system in which the distributions of properties do not change in space.

merits of a 'relativist' approach suggests that this concept has not yet been widely accepted. Many difficult questions in hydrogeology are related to the assumption of the REV. Such problems include upscaling, interpreting data from different types of measuring device and applying continuum approaches to systems.

It is worth considering at what point such averaging can reasonably be thought of as yielding continuum properties. The concept of the Equivalent Porous Medium (EPM) is often used to replace a fractured or channelled system with an equivalent uniform Darcian porous system. Berkowitz et al. (1988) provided a thorough examination of when an EPM approach is valid. Barker (1991) cited Sagar and Runchal (1982) who noted that a fracture system might be well characterised in terms of flow by an EPM model, but not in terms of transport.

EPM approaches are commonly applied to:

- The full porosity: **single EPM.**
- Two porosities of different conductivity: **dual permeability (DK).**
- Only the mobile porosity, coupled to an immobile zone: **dual porosity (DP).**
- Only the mobile porosity, coupled to immobile zones: **multiple porosity**

For all except the single EPM case, the exact position of the fractures relative to the blocks is not modelled, instead the systems are both regarded as overlapping continua.

For multiple porosity systems, once there is diffusive equilibrium between domains the single EPM may be a reasonable assumption (Berkowitz et al., 1988; Sanchez-Vila and Carrera, 2004). This can also be stated in terms of length scale; as long as the averaging scale is much larger than the fracture scale but much smaller than the domain scale, an REV can be reasonably applied to a fractured system. For a dense fracture network and short time the EPM may also be valid, except now characterising only the fracture porosity as there has been insufficient time for significant diffusion into the matrix (which is defined as the fraction of space in a porous medium that is relatively fine-grained, in contrast to channels or fractures which are 'open' and generally have increased local permeability). The possible fractal scaling behaviour of fractured rock might make this assumption highly questionable in some instances. Most commonly the ADE has been assumed to be the process which is applied to a single EPM.

Figure 3.2 shows how the upcoming parts of the thesis relate to this structure based around increasing heterogeneity and 'triangles' of processes. The next sections address each corner of each of the three 'triangles' in turn, examining the theoretical basis and literature for each. This starts with the simplest corner - **Advection Dispersion**.

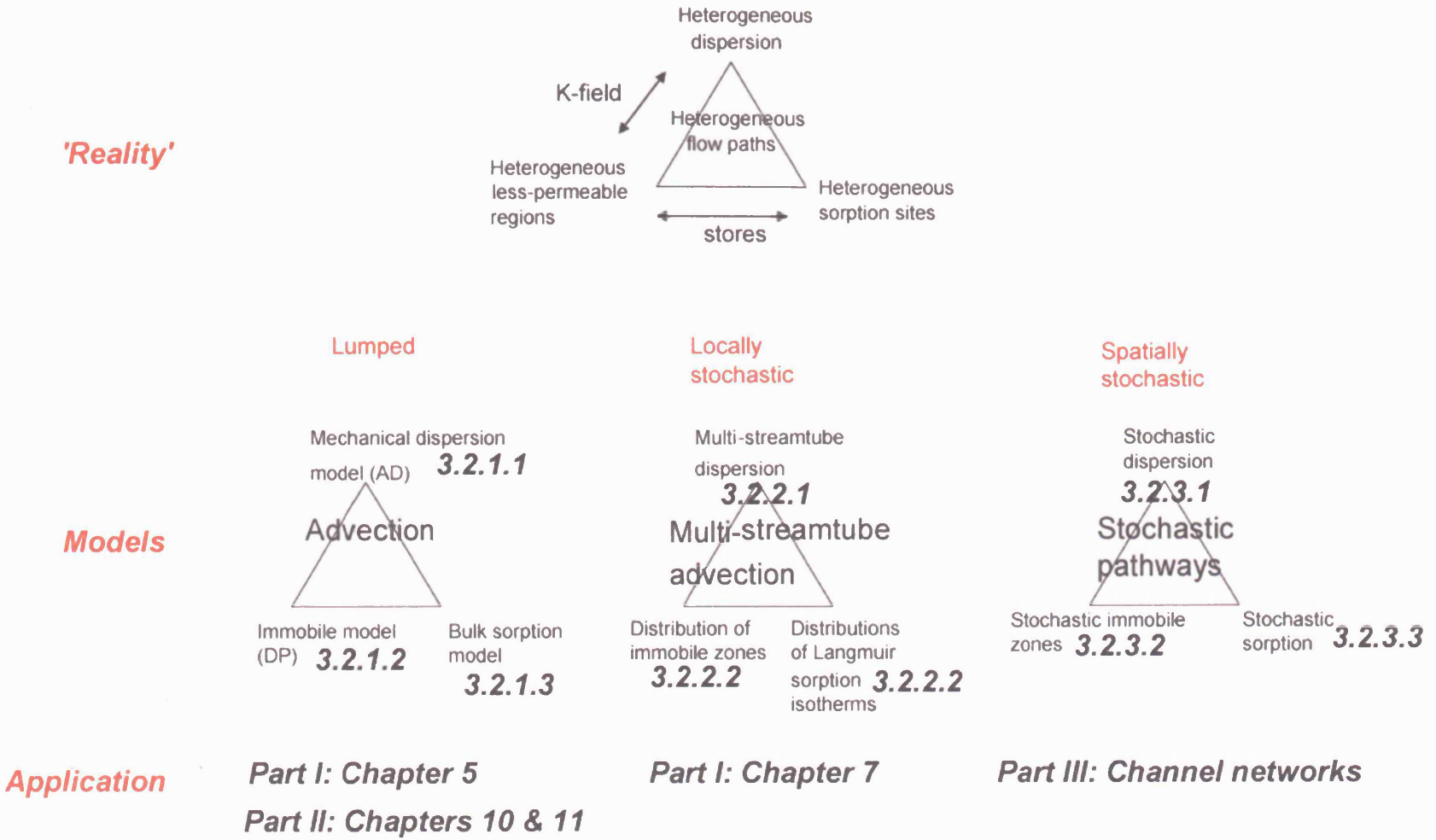


Figure 3.2: Incorporation of different models within the thesis (subsections and chapters where there is more detailed application indicated).

3.2 *Discussion of models*

3.2.1 'Lumped' models

3.2.1.1 The Advection-Dispersion process

3.2.1.1.1 Origins of the Advection Dispersion Equation (ADE)

The 1D ADE model (Equation (3.1)) is something of a 'benchmark' in transport modelling.

$$D\frac{\partial^2 C}{\partial z^2} - V\frac{\partial C}{\partial z} = \frac{\partial C}{\partial t} \tag{3.1}$$

where D is the (longitudinal) dispersion coefficient, C is the solute concentration and V is the average linear velocity.

There are several physically feasible conceptual models which can be shown mathematically to give rise to the ADE. Several approaches are illustrated in Bear (1972). This includes application of the Central Limit Theorem (CLT) to mixing within and translation between a series of identical cells (representing pores), to give a spatial concentration distribution that asymptotically approaches a Gaussian (Bear, 1960). The solution by Taylor (1953) relies on the lateral transfer of solute by diffusion across a parabolic velocity profile inside a cylindrical tube. A more general geometry is considered by Aris (1956). Roughness of channel or fracture walls can also produce mechanical dispersion at the channel scale and disparities in transport times between pathways within a network can produce dispersion at a yet larger scale.

The AD model is so ubiquitous that even basic definitions of dispersion (which could be due to many different processes) are standardised based on the AD model. The macrodispersion coefficient D_M is defined in terms of the second spatial moment,

$$D_M = \frac{1}{2}\frac{\partial \sigma_z^2}{\partial t} = \frac{V}{2}\frac{\partial \sigma_z^2}{\partial z} \tag{3.2}$$

Since in the instance of a unit impulse released between infinite upstream and downstream boundary conditions, $\sigma_z^2 = 2Dt$ and $D_M = D$, where σ_z^2 is the second spatial moment of concentration.

The ADE is an inadequate model for many field situations, (Carrera, 1993; Sanchez-Vila and Carrera, 2004). The spatial spreading of plumes are often observed to become non-Fickian (Berkowitz and Scher, 1998, 2001). Relatedly, in such instances, if the breakthrough curve (BTC) of a tracer is measured at a particular place and the ADE is fitted to the early part of the curve including the peak, the ADE often predicts a much sharper fall in concentration than the relatively long tail of the measured concentrations. This long tail is often attributed to diffusive

exchange taking place between mobile and less mobile zones, a condition under which the AD model may be inadequate.

Berkowitz and Scher (2001) and Berkowitz and Margolin (2004) showed that the ADE could be thought of as a special case of a continuous time random walk (CTRW). CTRW considers mass transfer at a small scale, by describing it in terms of random jumps between regions of different flow rate. Depending on the pdf of the random-jump, $\psi(s,t)$, this could predict AD or DP behaviour as sub-sets. The CTRW was fitted to the results from a laboratory experiment (Berkowitz and Scher, 1998) and fitted to the MADE breakthrough data. Berkowitz and Scher (2001:248) claim,

> "CTRW...can encompass a large number of physical processes with the appropriately determined $\psi(s,t)$, for example multiple trapping, multiple-rate models (e.g. Haggerty and Gorelick, 1995) and dispersion in stratified formations (Matheron and DeMarsily, 1980)"

The treatment by Carneiro (2000) showed that the method potentially has considerable practical applicability.

Some models have attempted to simulate heterogeneity at the domain (i.e. the total system volume enclosed by the outer boundaries of the system) scale, by varying dispersivity between large blocks within which the ADE applies for a given dispersivity (e.g. Bajracharya and Barry, 1997). Scheibe and Yabusaki (1998) illustrate a numerical modelling experiment that averaged K at different levels of resolution and showed that the transport predictions were a strong function of numerical cell volume.

On the assumption that dispersivity scales with the mixing scale, for field-scale macrodispersivity the mixed concentration is averaged over a large volume. At scales smaller than this averaging volume, there can be considerable local concentration variability (Cirpka and Kitanidis, 2000a & b; Kitanidis, 1994). For example, imagine two sharp impulses of tracer, one at $z = d$ and one $z = -d$, that therefore have a second moment of d^2. The tracer is spread out in one sense, but is clearly not as well dissolved as a bell-shaped distribution centred at $z = 0$ that has the same second moment (of d^2). This issue is important for reactive transport because it is *molecular-scale* mixing that controls reactions, not *macroscopic* mixing.

3.2.1.1.2 Theoretical values for dispersivity

Dominant dispersion regime

There are a number of possible scaling laws between D and V (Pickens and Grisak, 1981) of which a few of the simplest possibilities are highlighted here. This is potentially very important as it effects how solute will spread out depending on the velocity of flow. This could greatly affect flushing prediction, particularly if velocity is changed from that which may have been during a tracer experiment. The scaling of D and V for inter-grannular flow is shown in Figure 3.3.

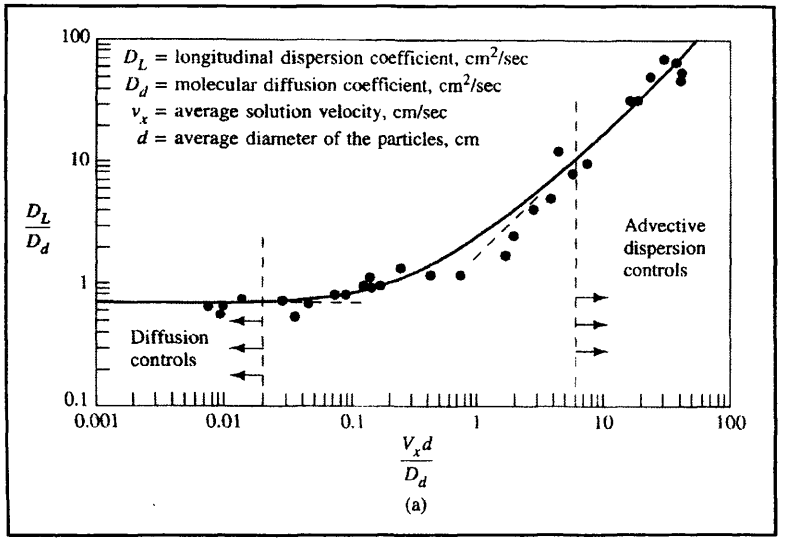


Figure 3.3: Dispersion coefficient in granular media (where V_x is V and D_d is D), from Fetter (1999:58), based on results from sand-column experiments by Perkins and Johnson (1963).

This data suggests the relationship $D = D_a + \alpha V$ (i.e. $D \approx \alpha V$ for $D_a \ll \alpha V$). Note that $Pe_g = Vd/D_a$ is a 'grain' Peclet Number, which is a measure of the ratio of time for diffusion across a grain (d^2 / D_a) to time for advection across a grain (d / V). For $Pe_g < 0.02$ transport is diffusion dominated (for example, potentially significant during flow through very low-permeability layers).

Detwiler et al. (2000) and Roux et al. (1998) worked on delineating the conditions for which the mechanical dispersion model is dominant. Detwiler et al. (2000) developed a condition for transition from a diffusion dominated regime to that of aperture-variation dominated dispersion, for a particular statistical model of fracture aperture variation. In this regime, $D \propto V$, as it is for uniform sand, when $Pe_g > 6$. Roux et al. (1998) examined scaling relationships for dispersion in an experiment where crack faces have been translated relative to one another. Dispersion due to roughness of the fracture walls can be superposed on to Taylor-Aris dispersion (see Bodin et al., 2003a).

Because parabolic flow profiles may remain locally valid in rough fractures, Taylor-Aris dispersion may apply to more realistic fracture geometries than idealised parallel plates (Bodin et al., 2003a). There is a second transition condition for yet higher V, for transition into this Taylor-Aris type dispersion regime.

Taylor-Aris dispersion asymptotically predicts $D = D^* + \frac{2}{105}\frac{a^2}{D^*}V^2$ for flow between parallel fractures and $D = D^* + \frac{1}{48}\frac{r^2}{D^*}V^2$ for circular channel flow (Bear, 1972) where a is the fracture half-width and r is the channel radius and D^* D_a is the coefficient of molecular diffusion. When diffusion is not dominating, Taylor-Aris predicts $D \propto V^2$.[11] This prediction for dispersion coefficient is an asymptotic result and does not occur until particles have had time to sample the whole of the velocity cross-section (which is parabolically-distributed across a conduit under laminar, Newtonian flow conditions) i.e. when $t \rightarrow a^2 / D^*$.

As velocities increase eventually turbulent flow ensues. It is also possible over longer distances that fracture or channel networks contribute to dispersion (see Figure 3.4).

[11] Note that away from the diffusion-dominated regime, D still scales with D^*, i.e. for the parallel fracture model when $\frac{2}{105}\frac{a^2}{D_a}V^2 \gg D^*$, then $D \propto 1/D^*$. This might lead to the possibility of confusion between an AD-dominated process and a DP-dominated process when examining BTCs for different tracers. Also note that further scaling laws will apply for dispersion where there is turbulent flow (for example, in a karstic conduit).

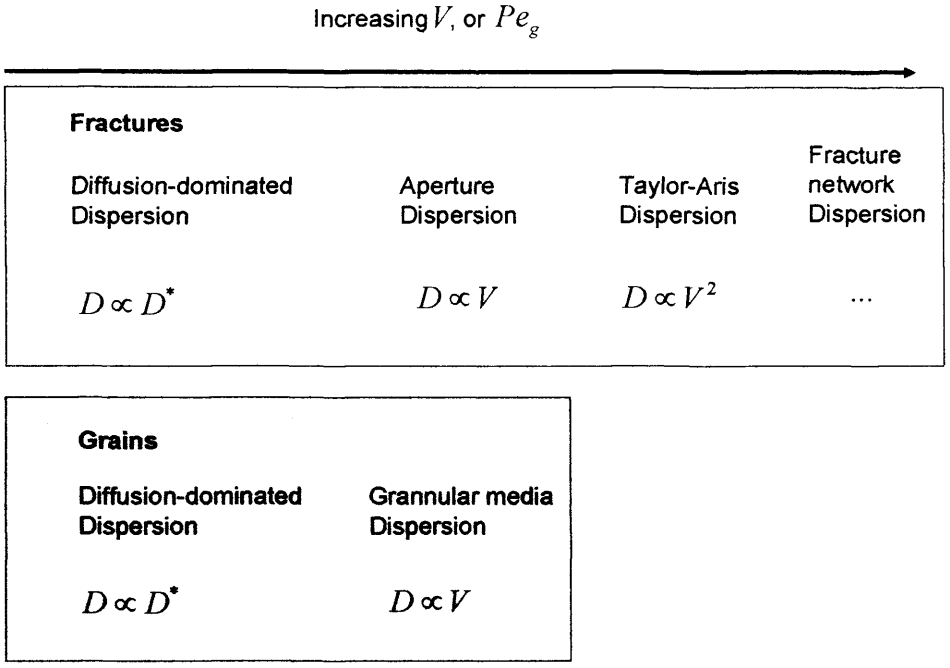


Figure 3.4: Simplified scaling relationships (sub network-scale).

One reason that these scaling relationships may potentially be important is that if an inappropriate relationship is selected for a given model this could contribute to incorrect process diagnosis and extrapolation in space and time may be significantly affected.

Scale effect

The most well-known scaling effect that has motivated a large body of stochastic hydrogeology literature is how dispersion appears to increase with distance. Through comparisons of many field experiments, Gelhar (1993) demonstrated that dispersivity apparently grows with distance (Figure 3.5). This 'scale effect' has been explained by noting that the tracer is likely to encounter heterogeneities at greater scales the further it travels. However, this is effectively saying that the REV does not apply. One focus of the stochastic hydrogeology literature has been to predict the evolution with distance of effective dispersion coefficients given different statistical descriptions of the aquifer. For example, Gelhar and Axness (1983) showed that for a lognormal K-field, the macrodispersivity tends asymptotically to a constant value after several integral scales.[12] Gelhar (1993) applied Taylor's approach to layered systems (diffusive exchange between layers, rather than across a parabolic profile). This analysis gives an initially linear growth in dispersivity that eventually tends asymptotically to a constant, but requires a band limited ('hole effect') spectrum of K values allocated to the layers. Matheron and de

[12] The integral scale is defined as $\lambda = \frac{1}{\sigma^2}\int_0^\infty R(z)dz$, where $R(z)$ is the covariance function and σ^2 is the variance (Gelhar, 1993).

Marsily (1980) also showed that by taking a velocity vector to be non-parallel to layering meant that the spectral stipulation could be dropped. Without the Taylor effect, the dispersivity of a stratified formation is predicted to grow linearly with distance. This is a prediction that only appears to hold true for short flow distances (where there has been little lateral dispersion); this is referred to as pre-asymptotic behaviour. Taylor theory might not hold for large scales as the assumption of lateral persistence of each layer is unlikely.

A number of studies have used a time-variable $D(t)$ (Dagan, 1988), or a distance-variable $D(z)$ in solutions of the ADE (Huang et al., 1996; Pickens 1981). There is a strong possibility that either the scale effect is a symptom of the AD equation being an incomplete description of the system or the scaling law for dispersivity is incorrect. The most useful evaluation of the AD model requires critical assessment of the latter issues and not empirical adjustment.

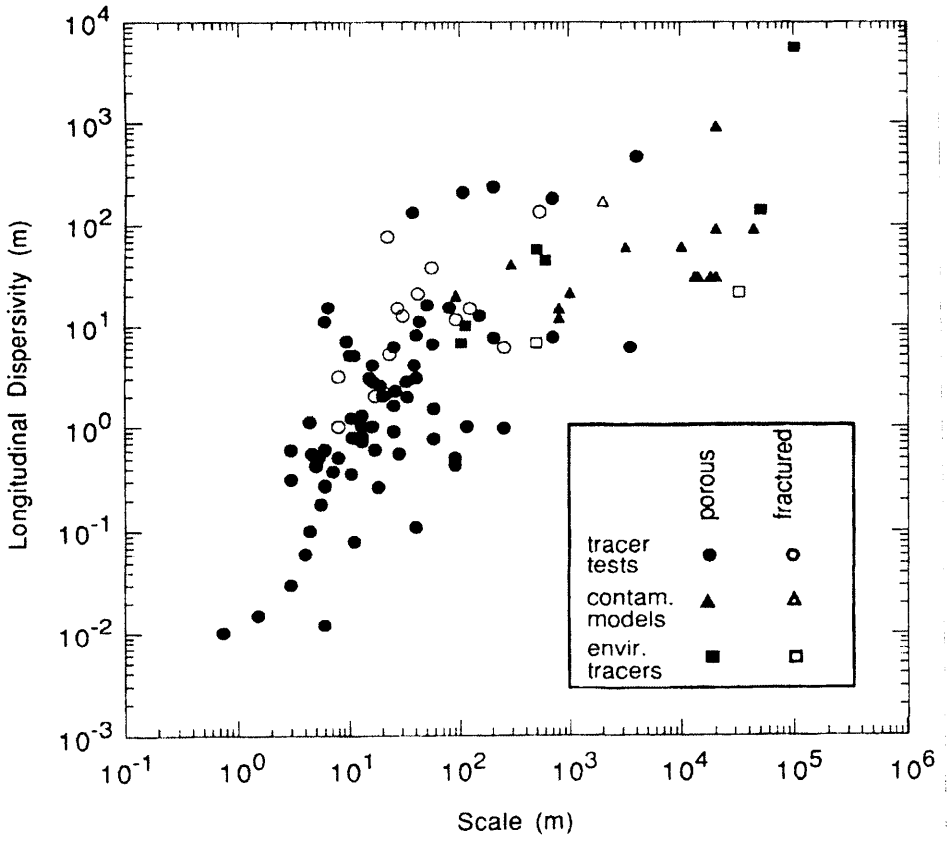


Figure 3.5: The scale effect (from Gelhar, 1993: 202).

A regression of the data in Figure 3.5 up to a scale of 1 km gives an approximate gradient of 0.1, which is the basis for the empirical relationship $\alpha = 0.1z$.

Flühler et al. (1996) identified a bundle of channels with no lateral mixing and channel velocities randomly selected from a single distribution, as a simple end-member dispersion

model. They noticed that it is the same as the Stochastic-Advective (SA) model of Jury and Roth (1990). Flühler et al. (1996) showed that the simple AD model is a second end-member and can be derived assuming complete lateral mixing of solute between such channels. They conceptualised solute mixing in soils as being bracketed between these two extremes. Javaux and Vanclooster (2003) and Zhang (2000) employed a hybrid model which is able to characterise the AD and SA processes as extremes and intermediate situations between them.

Sanchez-Vila and Carrera (2004) showed that the scale effect could not be explained by interpretation of what was really a DP process by using an AD model.[13] Many approaches guess 'reality' and then examine the scaling of the parameters of the AD fit.

Most simple analytical solutions and common numerical transport codes, such as MODFLOW's MT3DMS module or SUTRA, do not vary α with distance travelled, but require α to be fixed. Some caution should be expressed for spatial predictions close to the source if α has been set based on a far-away receptor. An example of this would be predictions of unrealistically large upstream dispersion from a source.

Barker[14] used a thought-experiment to establish an upper bound of $\alpha = 0.5\bar{z}$ by taking the extreme case of dispersion of two solute particles, one of which is stationary. This is remarkably close to the empirical result of $\alpha = 0.1\bar{z}$. One possible problem with this approach is that it is possible for the minimum velocity of a particle to be negative since some flow can be locally against the main flow[15]. Data for granular media rarely, if ever, shows greater dispersivities (e.g. Klotz and Moser, 1974).

3.2.1.2 Dual Porosity (DP) models

The term Dual Porosity (DP) normally applies to systems with bimodal permeability, where one region which is sufficiently more permeable than the other so as to render it 'mobile' whilst the other is considered to be 'immobile'[16]. These are often conceptualised as clumps or blocks of low K 'matrix' surrounded by high K 'pathways' or 'fractures'. The approximation is

[13] However, this was done using moments, which are not identical to comparing best-fits. Indeed, much of the stochastic literature is concerned with matching the second central moments. Questioning the wider validity of such generalisations partly motivates the method of multiple measures developed in the next chapter.

[14] In the Ineson Lecture, given by Prof. John Barker 28 October 2003, at the Geological Society of London.

[15] This might be the basis for the limit to dispersion for a dual-porosity model, where the minimum velocity would be zero when solute is diffused out of the mobile zone.

[16] The 'mobile-immobile' (MIM) model is here taken to mean a first-order dual porosity (DP) model and not taken to mean the same as a DP model which is taken to mean dual-porosity with diffusive exchange fully simulated.

commonly made because the conditions under which it might be reasonable include fractured rocks and aggregated soils. The simplest way of modelling the two regions in DP models is as overlapping continua, with the immobile blocks assumed to be in equal contact with the mobile region and having equal properties at all points (Baranblatt et al., 1960).

The nature of the immobile zone is important. It could be the boundary layer in the mobile zone (of say a fracture), pockets of open dead water (e.g. asperities on a fracture), dead-end channels or fractures, or else be contained within a fine porous matrix. The subscripts 'm' and 'im' for mobile and immobile water are used here to maintain a generalised discussion.[17]

The DP model may well have (partial) validity for approximating spatially stochastic systems and even for homogeneously packed systems through which flow self-organises into preferred routes. For such systems the point at which the approximation is reasonable is harder to define. For example Guswa and Freyberg (2002) gave a possible condition for macrodispersion (in a system with low permeability (K) lenses set in a homogenous material) of $K_{tot} \leq K_G$, otherwise they find formulations including mobile-immobile transfers are necessary.

Gelhar (1993) argued that real heterogeneous systems are better represented by a variable K field in which diffusional transport will be overwhelmed by advective effects. However, backed up by experiments with multiple diffusivity tracers, the DP concept has become well established under a range of circumstances (Maloszewski and Zuber, 1985, 1992, 1993).

The DP model category can be split into two: 'Diffusive' and 'Mobile Immobile' (MIM). It is easy to show for regular and known geometries that a 'diffusive' system tends towards behaving like a 'MIM' order system for long-time. Chapter 5 discusses DP models and Chapter 6 compares DP and MIM models.

3.2.1.2.1 Diffusive exchange

Fick's Second Law takes the form $\frac{\partial C_m}{\partial t} = \nabla.(D_a \nabla C_m)$. Assuming D_a is constant in space it becomes:

$$\frac{\partial C_m}{\partial t} = D_a \nabla^2 C_m \tag{3.3}$$

If it is assumed that the fracture concentration only changes significantly over a much larger distance than the block dimension, then the boundary condition at the block edge is the fracture

[17] The nomenclature for DP systems reflects what the two porosities are expected to be. Some authors use subscripts 'f' and 'm' to mean 'fracture' and 'matrix'.

concentration and can be considered to be constant all around the block. Classical analytical solutions to the heat equation for a step change in boundary conditions for simple geometrical shapes (such as rectangular parallepipeds, slabs, cylinders and spheres) are given in Carslaw and Jaeger (1959) and Crank (1975).

Barker (1985a & b) showed a general way of incorporating different geometries into the transport equation, by use of 'Block Geometry Functions' (BGF). One elegant aspect of this approach is that if the assumed geometry is changed, only the BGF need be changed, leaving the underlying form of the equation unchanged. The effect of different assumed geometries is discussed in several papers. For example, Zimmerman and Bodvarsson (1990) compared spherical, cylindrical and slab-like blocks in an unsaturated flow context.

Analytical solutions to the transport equation assuming DP exchange are rare and difficult to implement (e.g. Skopp and Warrick, 1974; Sudicky and Frind, 1982). However, there are numerous ways to solve the diffusion equation within a transport code. Each way is appropriate to different types of problems with the principle dictated by computational time, that simpler models are more appropriate for simpler problems. Many more complex approaches exist[18], however only the simplest conditions are adopted in this work:

(i) Take only first term[19] (e.g. the MIM exchange in the Hydrus model (Šimunek et al., 2005)).

(ii) Truncate the infinite series by taking a number of terms for a given accuracy (e.g. Bibby, 1981)

(iii) Laplace transform (LT) inversion, analytically or numerically.

(iv) Semi-analytical approaches (e.g. DP-PULSE, (Barker et al., 2000)) that can cope with piecewise constant C or V.

[18] More complex conditions can be analysed by techniques such as:

(i) Strings of cells away from fracture, to simulate diffusion by finite difference (e.g. MINC, (Pruess and Narasimhan, 1985)).

(ii) Galerkin finite elements combined with LT inversion (Sudicky, 1990a&b; Sudicky and McLaren, 1992a &b).

(iii) Application of Herrera's memory functions (Herrera and Rodarte, 1973; Herrera, 1974; Herrera and Yates, 1977; Premchitt, 1981) to DP allows a very efficient recursive solution, that does not require solving a convolution integral for all time. Barker (1985) first noticed the potential and Carrera et al. (1998) later developed it into a working model called TRANSIN. See also Little et al. (1996).

(iv) Empirical matrix exchange term to avoid an infinite series without a large divergence from the diffusive model. (e.g. matching the MIM expression in some way (Zimmerman and Bodvarsson, 1989; Zimmerman et al., 1993)), assuming a linear pressure profile a certain depth inside the boundary (Dykhuizen, 1990), or the 'trial function' method of Vinsome & Westerveld (cited by Pruess and Wu, 1993).

[19] This thereby becomes a MIM expression (see next section).

3.2.1.2.2 Mobile-Immobile (MIM) exchange

In contrast to Fick's Second Law, a simple expression for mass transfer is given by the MIM:

$$\theta_{im}\frac{\partial\langle C_{im}\rangle}{\partial t}=\kappa\left(C_m-\langle C_{im}\rangle\right) \tag{3.4}$$

where the rate of transfer is proportional to the difference in the mobile and average immobile concentrations (C_m and $\langle C_{im}\rangle$) and the exchange coefficient, κ. This concept is variously described as 'Mobile-Immobile' (MIM), 'first-order', 'dead-end pore', 'Two Region Model' (TRM) and 'Quasi-Steady State' (QSS).

Expressing this in dimensionless time:

$$\frac{\partial\langle C_{im}\rangle}{\partial T}=\frac{\alpha t_{cb}}{\theta_{im}}(C_m-\langle C_{im}\rangle)=\kappa'(C_m-\langle C_{im}\rangle) \tag{3.5}$$

where $\kappa'=\kappa t_{cb}/\theta_{im}$, $T=\frac{t}{t_{cb}}$ and $\langle C_{im}\rangle$ is the concentration averaged throughout the immobile zone. The meaning that should be attached to κ (or κ') is connected to whatever justification is used for use of this model. If it is intended to be a purely empirical representation, then it forms little more than a fitting constant. If it is intended to provide an approximation to a diffusion model with a known geometry then κ is related to the parameters of the diffusive model in a number of ways.

The MIM model has the advantage of being simple, efficient to solve and is in many cases, commensurate with the level of ignorance of the real geometry. If the geometry is highly complex and possibly even unknowable, then it may be no more arbitrary to assume a MIM transfer than a diffusive one (Nkedi-Kizza et al., 1983).

It is conceivable that in the instance of channel flow next to relatively immobile water, that non-diffusive mass transfer takes place by means of eddies (Raven et al., 1988) or other flow structures. In such a case the diffusion equation would be physically inappropriate and a MIM model might be adequate. First order kinetic sorption might also be modelled with the MIM. Thus the MIM has the strength of providing a 'grey box' model, without necessarily needing the constituent processes or geometries to be deterministically examined. However, the MIM model is unlikely to ever correctly represent the underlying physics, so requires careful treatment as a useful approximate model.

3.2.1.3 Sorption

The third category of lumped models is bulk sorption. There is a substantial literature devoted to the chemistry of surfaces giving rise to different isotherms. In the framework used here, the very simplest isotherms are applied to the basic continuum case, in keeping with having the basic 'benchmark' AD and DP models also at this level. As is the case with the AD and DP models, the most basic processes appear in standard texts. For sorption this includes Fetter (1999), Freeze and Cherry (1979) and Appelo and Postma (1994).

Consider a monotonic isotherm $C^* = \mathrm{f}(C)$, where C^* is the sorbed mass per unit mass of dry solid. Transport in the presence of equilibrium partitioning with this isotherm can be simplified in the absence of any exchange processes, into an ADE with sorption to give:

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - V\frac{\partial C}{\partial z} - \frac{B_d}{\theta}\frac{\partial C^*}{\partial t} \tag{3.6}$$

where B_d is bulk density.

Using the chain rule to differentiate the isotherm, gives:

$$\frac{\partial C^*}{\partial t} = \frac{d\mathrm{f}(C)}{dC}\frac{\partial C}{\partial t} = \mathrm{f}'(C)\frac{\partial C}{\partial t} \tag{3.7}$$

Substituting back into Equation (3.6) gives:

$$\frac{\partial C}{\partial t}\left(1 + \frac{B_d}{\theta}\frac{dC^*}{dC}\right) = D\frac{\partial^2 C}{\partial z^2} - V\frac{\partial C}{\partial z} \tag{3.8}$$

Define a retardation coefficient, $R = 1 + \frac{B_d}{\theta}\frac{dC^*}{dC}$ gives,

$$R\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial z^2} - V\frac{\partial C}{\partial z} \tag{3.9}$$

R depends on the relationship between $C*$ and C, i.e. on the shape of $\mathrm{f}(C)$.

Three basic models considered here are the linear, Freundlich and Langmuir isotherms.

Linear: $C* = K_d C$, $R = 1 + \frac{B_d K_d}{\theta}$, where K_d is the distribution coefficient.

Freundlich: $C* = KC^N$, $R = 1 + \frac{B_d KNC^{N-1}}{\theta}$, where N is the Freundlich exponent.

Langmuir: $\frac{C^*}{C} = \frac{\alpha\beta}{1+C\alpha}$, $R = 1 + \frac{B_d\alpha\beta}{\theta(1+\alpha C)^2}$, where α is an adsorption constant and β is the maximum amount of solute that can be adsorbed.

These three forms of isotherm can potentially be confused in the presence of noise: possibly a linear fit could be made inappropriately to either non-linear curve, or a Freundlich fitted to a truly Langmuir system or vice versa. It would be dangerous to extrapolate any of these isotherms beyond the data. One example would be fitting a linear isotherm and then modelling at higher concentrations, where actually the isotherm is curving. Alternatively the model might be extrapolated for small C, below that measured on the isotherm.

Measurements of real isotherms can reveal hysteresis, exhibiting different sorption 'rising limbs' to desorption 'falling limbs'. Appelo and Postma (1994) explained that often this is an apparent effect, perhaps a symptom of diffusive non-equilibrium, losses due to reaction or accumulating errors. There follows a few points on each of the three simple isotherms in order to introduce them in more detail.

(i) Linear isotherm

The linear isotherm is especially convenient for transport models as it only introduces a constant into the ADE that tends to just delay the time; thereby keeping the equation linear. The obvious change is that all times, t, are replaced by t/R. The same concentration is therefore reached at the later time of tR. As an example of linear retardation, consider the retarded step AD solution (Ogata and Banks, 1961):

$$\frac{C}{C_0} = \frac{1}{2}\operatorname{erfc}\left(\frac{z - Vt/R}{\sqrt{4Dt/R}}\right) + \frac{1}{2}e^{(Vz/D)}\operatorname{erfc}\left(\frac{z + Vt/R}{\sqrt{4Dt/R}}\right) \tag{3.10}$$

(ii) Freundlich isotherm

The Freundlich isotherm has only two parameters and has proved successful for modelling many real datasets. It is conceivable that certain materials have distributions of surface energies that give rise to a Freundlich isotherm (Hayward and Trapnell, 1964; Kinniburgh et al., 1983). The Freundlich isotherm allows indefinite increase of the mass of sorbed solute, which is physically not possible. However, the 'modified Freundlich' equation (Kinniburgh, 1986) is an attempt to remove this problem for high concentrations by normalising C by the solubility.

The Freundlich isotherm does not tend to a linear relationship for low concentration. The question of whether this is physical or not can only be answered by reference to specific datasets. The linear isotherm is, however, a subset in the instance that the Freundlich exponent N is equal to 1. For other values of N the isotherm will be either concave or convex.

(iii) Langmuir isotherm

The Langmuir isotherm can be used as a building block from which other isotherms may be constructed. The basic isotherm is easy to derive using a simple kinetic argument and a finite number of sites. Derivations of the Langmuir isotherm are given by Hayward and Trapnell (1964) using kinetic, statistical and thermodynamic arguments.

This isotherm is inherently convex and tends to linearise at small C. [20] At high concentration, C^* approaches a physically reasonable maximum.

It is the more minimally parameterised Freundlich isotherm which is explored in this thesis. Langmuir sorption is used in Chapter 7 as a basis for building up different representations of heterogeneous sorption. This concludes the introduction to sorption, which was the final process in the least hetereogeneous group in Figure 3.1 - 'lumped models'. The next heterogeneity level is now examined.

3.2.2 'Locally stochastic' models

A simple way of representing heterogeneity that gives more detailed behaviour than the simple lumped models is now introduced. This level is given the title 'locally stochastic', to convey the concept that properties are given simple distributions. This will be seen to be in contrast with the third heterogeneity level, that allows properties to be characterised by spatially stochastically variables. The same process triangle as used for the lumped models in Figure 3.1 is now applied, but now considering distributions of lumped processes.

3.2.2.1 Multi-stream-tube advection and dispersion

Firstly, the dispersive process can be pictured as the sum of many advective processes. A multiple stream-tube approach considers transport through more than one stream-tube.[21]

---

[20] **Physically** a Langmuir-type relationship might be expected at low concentrations when the sorbing molecules are unlikely to interact one another or affect the other sorbtion sites. Even more simply, it might be expected that the rate leaving is proportional to the number of occupied sites (i.e. $\propto \theta$) and the rate joining is only proportional to the concentration of free molecules (i.e. $\propto C$). In other words, the number of free sites is no longer significant. Immediately, the balance becomes $\theta = \alpha C$ or $C^* = \alpha\beta C = K_d C$, which is linear. **Mathematically** the Langmuir isotherm for small C tends to $C^* = \alpha\beta C = K_d C$, since the denominator tends to 1. A two-site Langmuir tends to a linear isotherm for small C, since $\frac{C^*}{C} = \frac{\alpha_1\beta_1}{1+\alpha_1 C} + \frac{\alpha_2\beta_2}{1+\alpha_2 C}$ tends to $C^*/C = \alpha_1\beta_1 + \alpha_2\beta_2$ or $C^* = K_d C$. Indeed from this it can be seen that any (finite) number of Langmuir sites would add up to give $\alpha_1\beta_1 + \alpha_2\beta_2 + ...\alpha_N\beta_N = K_d$.

[21] A stream-tube is defined as a closed volume bounded by inlet and outlet surfaces and streamlines. It can be thought of as a bundle of streamlines (Massey, 1989).

Stream-tubes could represent non-interacting discrete channels, but also could represent hydrostratigraphic layers, or individual fractures or high-K pathways through a stochastic continuum.[22]

The most basic application is a set of parallel stream-tubes which only advect ('multi stream-tube advection' or the 'SA' model). Different distributions of (piston flow) travel times give rise to different predictions (for example, a lognormal distribution of times (Jury and Roth, 1990)).

A more complex application is a set of parallel stream-tubes which advect and disperse (i.e. a set of AD tubes, perhaps representing several discrete channels). Dual permeability (DK) models represent the very simplest example whereby there are only two unconnected stream-tubes (for example the 'multi-dispersion model' of Maloszewski and Zuber (1992) for use in karstic systems which is essentially the superposition of two AD models). These stream-tubes could be coupled by an exchange term. This gives rise to a model similar to dual-porosity models (which are discussed in Section 3.2.2.2), but there need not be any stipulation about the relative values of K in the different domains. For unsaturated zone flow there are several models that adopt a DK approach, with a 'fast-flow' component for flow down fractures, or 'macro-pores', coupled to a slower flowing 'matrix' (e.g. the MACRO model (Javis, 1991)). DK models can produce BTCs that are difficult to distinguish from DP models. However if the characteristic transport times of the two systems are significantly different and the spreading due to transfer and mixing within each region is small enough, they can also yield conditions in which distinguishable double peaks are created (Ma and Selim, 1995). Gerke and van Genutchen (1993) and van Genuchten (1996) presented a DK model. They solved Richard's equation in two domains, that have a water flux between them based on a first order exchange term. There is also a diffusive flux between the two. The total solute transferred between the two is the sum of the diffusive and advective fluxes.

So, by assuming or finding a distribution and making a simple summation or integral, the basic advective or AD models can be combined to give a broader variety of locally stochastic models.

[22] This approach has received much attention, especially for reactive transport (Bardsley, 2003; Cirpka and Kitanidis, 2000; Flühler et al., 1996; Jury and Roth, 1990; Neretnieks, 2002; Toride and Leij, 1996a& b). The Lagrangian stochastic branch of stochastic modelling describes $V(t)$ statistically as the particle moves along the stream-tube (Cvetkovic et al., 1992). This has been developed in a series of papers considering reactive transport (Cvetkovic et al., 1992; Cvetkovic and Dagan, 1994; Cvetkovic et al., 1998; Dagan and Cvetkovic 1996).

3.2.2.2 Multi-store models (DP and sorption)

Locally stochastic approaches can also be applied to immobile regions. Multiple porosity models have one mobile region and a number of immobile regions that relate to each other in different ways. This could relate to a series of different stores directly connected to the mobile region or possibly nested in a hierarchical fashion. Bai et al. (1993) discussed the general concepts of multiporosity and multipermeability.

Three-region models (triple-porosity models) are the next step in complexity from DP models. It might be that the pores of the immobile zone at the edge of an aggregate, or matrix block may be larger and more accessible than deeper into the immobile zone perhaps due to physical or chemical damage. There is thus the case for modelling a 'fast' and a 'slow' immobile porosity. Gwo et al. (1995) developed a three-region model for soils, made up of micro-, meso- and macro- pores. Wu et al. (2004) developed a triple continuum model for fractured rock, justifying the model on the basis that there was a main set of large fractures in addition to a second smaller fracture set, which was in contact with the matrix. In a fractured system there is a strong chance due to variable apertures and channelling that there will be immobile water in the fractures, as well as immobile water in the matrix. Bodin et al. (2003b) advocated a three region model to simulate this effect.

It is simpler if the transport model can be cast without requiring a transient flow model. In reality, advective transport from a block occurs if pressure gradients are set up between the matrix and fracture, for example following the commencement of pumping. This could be approximately simulated by having an extra 'fast' store which responds more rapidly, thereby removing the need to solve the unsteady pressure equation.

The transfer in a three-region model could in principle be diffusive. It is conceivable that the diffusion equation could be solved for a geometry that included a finite thickness outer skin, but it might be difficult to obtain physical evidence to support the skin thickness assumption. Barker (1985b), de Swaan (1990), Moench (1984) and Moench, 1995 have all invoked fracture skins.

More commonly, a three region model has a MIM formulation. Different permutations of coupling will depend on the conceptual model:

- All three regions coupled (as per Wu et al., 2004).
- Both immobile regions coupled to mobile region but not each other.
- Mobile region coupled to immobile region 1, which is in turn coupled to immobile region 2.

The three-region model is pulled out here because like the two-region model it is a common special case. Thereafter any number of regions could be imagined, or even a continuous distribution of regions.

Haggerty and Gorelick (1995) showed that a specially chosen distribution of an infinite number of first order rates (MIM exchanges) yield a model that identically mimics a diffusive DP model with a simple geometry. In the same paper they also showed that many first-order physical and chemical models that appear in the hydrogeology and soils literatures are mathematically identical. This impressive result is re-demonstrated here, giving good results even for fairly low numbers of terms in the sum. This is virtually the same as the demonstration of accuracy by Bibby (1981) of a truncated sum for the diffusion solution.

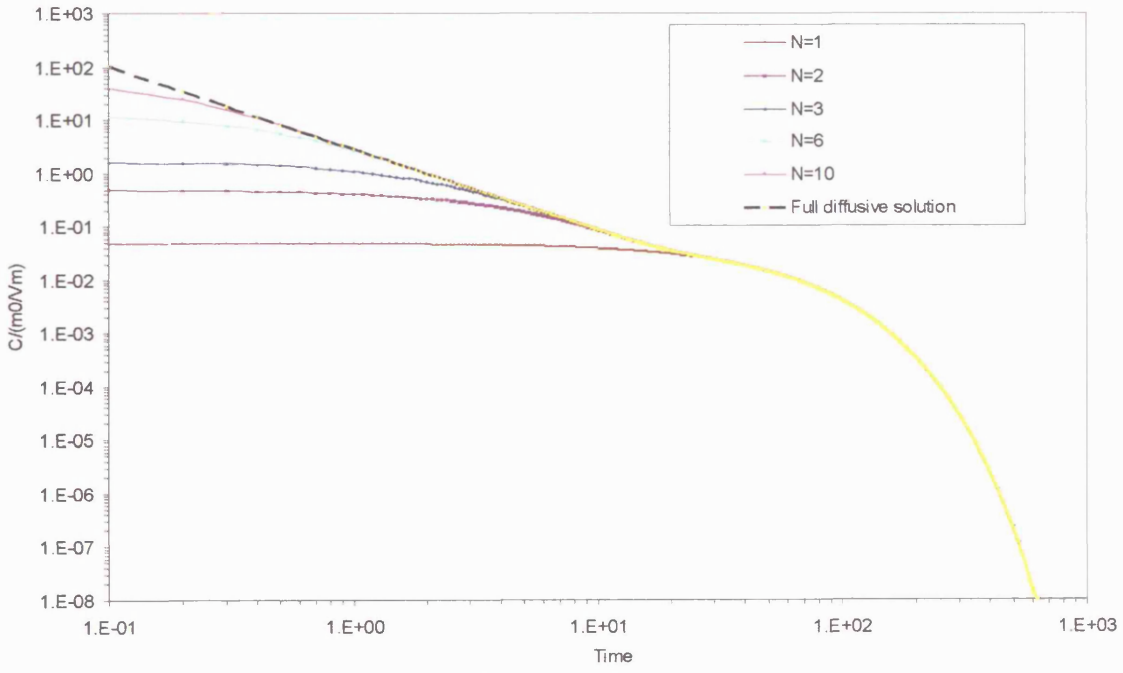


Figure 3.6: Demonstration of how few correctly-chosen MIM terms are needed to mimic the diffusive impulse solution (slab) for all time except the very earliest. $t_{cb} = 100$, $t_a = 0.01$, $\sigma = 1$ and $m_0 / V_m = 10000$.

If the fully diffusive model is being used then there are various ways of looking at heterogeneity. Instead of a single characteristic exchange timescale, t_{cb}, there could be considered to be a distribution of characteristic timescales. The components of t_{cb}, as related by $t_{cb} = b^2 / RD_a$ (where b is block volume/area, R is the retardation coefficient and D_a is

the apparent diffusion coefficient), may themselves be variable. In MIM terminology this could be a distribution of $t_M = \theta_{im} / R\kappa$.

A further way of conceptualising the distribution of scale, b, is to consider the distribution of the volume, V, occupied by a given scale. Barker (1985a&b) described how a composite BGF could be made up as the sum of N separate blocks, with discrete probability, P_i, or as the integral of an infinite number of blocks with continuous probability distribution, P. Barker's method was originally based on volume distributions, although it could be generalised to any distribution (see Chapter 7). The MINC model (Pruess and Narasimhan, 1982) in principle allows for variable block sizes and shapes by use of 'proximity functions'. Van Genuchten and Dalton (1986) used a 'shape factor' approach which amounted to a similar approach to Barker (1985a). Zimmerman and Bodvarsson (1990) introduced a scaling law for irregularly shaped blocks (e.g. polygonal shapes formed by the intersection of non parallel fracture sets) based on the ratio of surface area to volume. D_a will vary due to differences in tortuosity, pore size and connectivity.[23] Different tracers may be able to access different porosities. Anionic exclusion, for example, is a well documented mechanism in the presence of clay minerals (Yu and Neretnieks, 1997).

Haggerty (1999) puts all these effects together suggesting there is a good case for simultaneous variability of b, D_a, pore geometry (shape, tortuosity, cross-sectional area and connectivity) and R.[24] Heterogeneous kinetic sorption may also potentially contribute an indistinguishable additional heterogeneity of times. The combined effect of all the variability of all these parameters is potentially a wide distribution in characteristic times.

3.2.2.3 Heterogeneous Sorption

Analogous to distributions of DP block sizes, heterogeneous sorption can be conceptualised as a distribution of surface sites. This is detailed in Chapter 7.

Despite Freundlich having first postulated this isotherm as an empirical relationship, various later studies showed that it might have a theoretical basis as the result of a particular type of heterogeneity. Hayward and Trapnell (1964) showed how an integral over a distribution of Langmuir isotherms (by an exponential distribution of heat of adsorptions) yielded a Freundlich

[23] Ohlsson and Neretnieks (1995) documented two orders of magnitude differences in diffusivity in granitic rocks at Äspö hard rock laboratory.

[24] In many instances it may be impossible to separate the variability in b, D_a and R.

isotherm. This is not universally acknowledged in the literature (Hinz et al. (1994), for example, stated that there is no such distribution).

Hayward and Trapnell (1964) demonstrated that an increase in repulsion with occupancy is an insufficient condition to generate a Freundlich isotherm. In other words, aside from a good empirical fit, Hayward and Trapnell (1964) found that the only viable underlying physical explanation would be heterogeneity. Speculatively, perhaps some systems that generate a Freundlich isotherm might continue to give Freundlich-type behaviour at concentrations below what was measured, as the underlying distribution of sites is important. None of the data presented by Kinniburgh (1986) does this however. Without an argument to disprove this possibility, in the absence of isotherm data down to the desired concentration, caution should be applied in linearising isotherms to model low concentration transport, unless the data has already become unambiguously linear.

Statistical thermodynamic approaches assume that the heterogeneity is at the level of the sites. It is possible that isotherms are built up due to the homogenisation of heterogeneous materials exhibiting heterogeneity at a much larger scale than between individual sites, but how a Freundlich isotherm would result in this instance is unclear.

Although theory cannot completely guide the best way to proceed, the locally stochastic method offers great flexibility and gives the chance to test a wider range of conceptual possibilities than are embodied in lumped sorption models.

3.2.3 'Spatially stochastic'

The third level on the heterogeneity axis of Figure 3.1 is 'spatially stochastic' models. It is best to think of these as the 'most heterogeneous' representation, but not necessarily the 'most realistic' or 'most complex'. This level of representation has had a lot of literature attention, despite the paucity of data to inform it. The attention given to this level throughout the thesis is commensurate with its applicability to waste data (i.e. it is potentially useful and worth speculative investigation, but current data are far from being able to constrain such models).

3.2.3.1 Continuous variability in K

The flow domain can be represented as a continuum in which K is distributed according to different stochastic representations. If the integral scale of the highest and lowest K values are relatively long, there is the possibility of representing fractures and flow barriers. A bi-modal permeability field to represent a fractured rock is one sub-set of possibilities available to this

approach. These types of model, because of their generality, could occupy a substantial proportion of 'model space'. A potential weakness is that geostatistical data may be insufficient to properly support the assumed models.

Summaries of stochastic models are given by Dagan (1989), De Marsily (1986), Gelhar 1993 and Jury and Roth (1990). Much of this work has focussed on relating aspects of the spatially stochastic distribution (especially σ_Y) to mean values of macrodispersion coefficient and possible predictions of scaling of this coefficient.

As an alternative way of modelling media that are sometimes conceptualised as DP, Tsang et al. (1996) produced fields of stochastically generated K (with long-range correlation structures given to the extreme Ks). It is not surprising that this conceptualisation might give rise to DP-type behaviour, as Carrera et al. (1998) noted that a stochastic formulation of the groundwater flow equation with spatially variable Darcy velocity, q, gives a term that looks like a source-sink term that is due to heterogeneity, rather than matrix transfer. Tsang et al. (1996) used ISIM3D (Gomez-Hernandez and Srivastava, 1990) to generate a stochastic K-field based on a measured log-K distribution, rather than fitting an assumed distribution, such as a lognormal. They found that spatially integrated (lumped) predictions are more appropriate than predictions of concentrations at specific points, given the resolution of normal datasets. Carrera (1993) also aligned with this view and to a great extent it is the stance of this thesis.[25]

Zinn and Harvey (2003) found that a univariate distribution of K and a spatial correlation function may be inadequate to describe transport. They note that it is possible to have different levels of connectedness (of extreme high and low permeabilities) with the same overall spatial statistics, but with very different transport behaviour. They found another rough categorisation of field conditions. For a multi-gaussian K-field, but with high K values well connected they found:

- A low K contrast (σ_Y^2<4) gave little mass transfer and the ADE was adequate.
- A medium K contrast (4<σ_Y^2<8) had advection through the low-K regions dominating the tailing.
- A high K contrast (σ_Y^2>8) showed both diffusion and advection to be important (classic DP behaviour).

[25] A greater weighting is given to lumped models in this thesis, although it is certainly the case that stochastic representations are both useful and important. Part III of this thesis addresses one such stochastic model, which potentially gives very useful insights about the behaviour of sparse channel systems, but it would be difficult to calibrate.

This serves as a warning that bulk statistics can miss subtleties which turn out to dominate transport.

3.2.3.2 Bimodal K (an important subset of spatially stochastic models)

An important subset of spatially stochastic models is the instance of two distinct K values that gives a flowing system and a relatively immobile one. Two primary categories of models arise in this instance - discrete fracture networks and channel networks.

Discrete Fracture networks

In a review paper, Berkowitz (2002) described how critical the details of fracture geometry are for many fractured rocks, yet how poorly resolved even the best datasets are.

One way to handle data uncertainty in fracture geometry is to model probabilistically by generating ensembles of many realisations selected from pdfs of parameter values. However, the pdfs themselves are subject to considerably uncertainty, in particular the statistics of fracture aperture distribution is not very reliable (Bodin et al., 2003b).

Variable length scales mean discrete fracture models span a spectrum from densely fractured systems that can be idealised as EPM, through to single fracture models.

An intermediate category of discrete fracture networks exist, in which the matrix is a permeable EPM. This is justified for instances where there are a small number of dominant fractures which need to be modelled explicitly and where smaller scale fractures can be lumped together and idealised as an EPM. Carrera and Martinez-Landa (2000) described use of such models.

Dense DFN models (e.g. Fracman, Dershowitz et al. (1995)) may require upscaling to an EPM model. This is an ongoing challenge, however there are various tools that attempt this, such as Fracman's StrataFrac module for upscaling porosity and permeability.[26]

Channel networks

There is evidence that the flow along fracture planes tends to concentrate in particular preferential paths, because flow through fractures is very sensitive to local variations in aperture. There is both modelling evidence (Tsang et al., 1991) and field evidence (Abelin et al., 1991) to support the hypothesis of channelling at fracture intersections. It is possible that networks may arise in more general settings than in fractured rock due to interconnecting preferred paths or channels.

[26] http://fracman.golder.com/software/stratafrac.asp.

In channel systems there is a reduced surface area of mobile water in contact with the immobile zone. Barker (1985a) and Young and Ball (1997) gave an expression for diffusion from a circular sectioned channel into a surrounding cylinder of matrix.

More simply, Cornaton and Perrochet (2002) applied a MIM formulation (with somewhat limited success) to a simple branching karst network, conceptualising the non-channelled finely fractured limestone as the 'matrix' which exchanges with different channels.

There are explicit channel network models in existence, for example Tsang et al. (1987) and the CHAN3D model of Moreno and Neretnieks (1992). CHAN3D was used by Gylling et al. (1998) to model tracer breakthrough at Äspö. Channel conductance for each channel was selected independently from a lognormal distribution (i.e. assuming spatially uncorrelated conductances). The model included a diffusion exchange term between the channels and the matrix.

For sparse networks connectivity is of crucial importance. Percolation theory (Berkowitz and Balberg, 1993; Hunt, 2005; Sahimi, 1994) predicts the scaling of properties depending on the level of connectivity - an observation with potentially powerful predictive implications.

Jourde et al. (2002) presented a channel network flow model for the case in which different sized channels form at different joints ('T' and 'X' joints) in a rock with two main sets of fractures. They compared the drawdown for a generalised radial flow model (Barker, 1988) with a transient pipe-network solution and found that the hydrodynamic response of both models is similar. Although this was a flow model the lesson is that it may be possible to misdiagnose pipe-flow as a 'classical' DP system.

Chapter 13 of the thesis examines channel networks as an example of the second type of bimodal K models. A code called HyperConv (Robinson, 2003) is used to investigate flow properties of fracture networks in Part IV. It is important to examine such models even in the absence of data, because they may give insights by exhibiting emergent properties that are similar to reality and because new models are important to continue to stretch the envelope of prediction, thus avoiding non-conservative estimates based on a limited and unimaginative model set.

3.2.3.3 Spatially stochastic sorption

The final model category in Figure 3.1 is spatially variable sorption.

Cvetkovic and Dagan (1994) and Dagan et al. (1992) developed a Lagrangian nomenclature for general reactive transport though heterogeneous media. In terms of relating to bulk properties they focus on developing integral expressions for spatial and temporal moments. Cvetkovic et al. (1998) applied this approach to systems that have non-linear sorption (Freundlich specifically). Berglund and Cvetkovic (1996) built on this discussing flushing and late-time tailing and pointing out that non-linear sorption can considerably prolong aquifer clean up times.

Rabideau and Miller (1994) modelled spatially variable sorption properties as an additional property to a model that already coped with spatially variable K. They assigned a lognormal distribution of K_d (selecting the same variance as was selected for the lognormal distribution of K) in both an uncorrelated way and a correlated way, with the same correlation length for K as for all cells and found a greatly longer flushing time when there was a strong negative correlation.

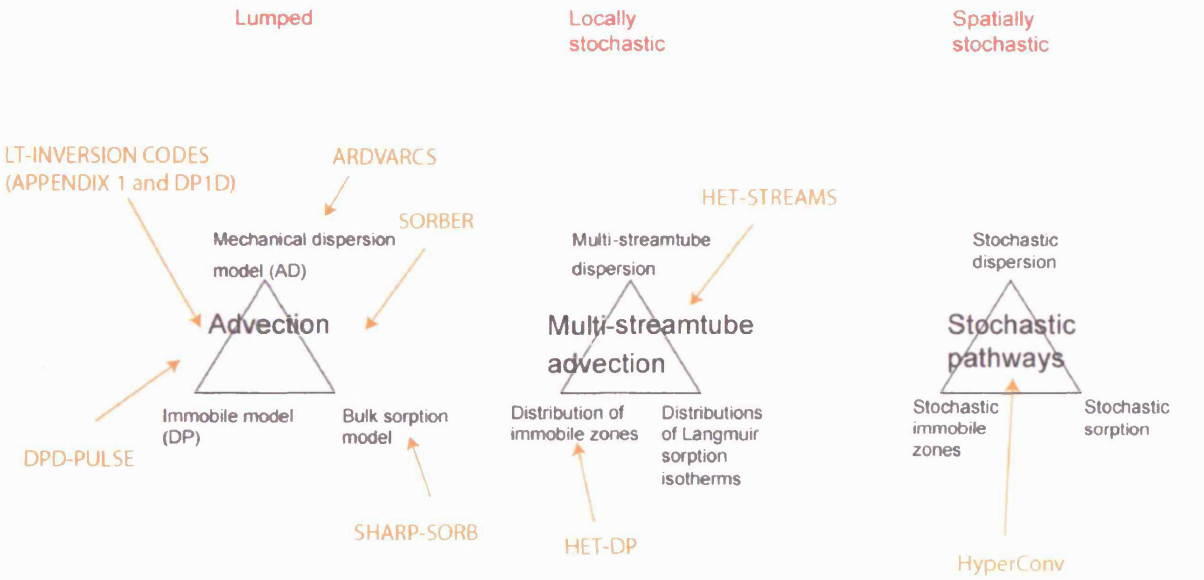
The difficulty with using this class of models is obtaining data with which to constrain models, but it is still useful in terms of finding out how important this effect might be and also in informing possible data collection requirements.

3.3 *Summary*

In order to predict transport in highly heterogeneous media, appropriate models must be selected. As a contribution towards this, a framework of mass-transport models has been suggested in this chapter (Figure 3.1) which allows broad categorisation of models in which heterogeneity is the primary axis.

Where there is conceptual uncertainty, 'diagnosis' of different models may be required based on BTC or flushing data. If a particular model or model group cannot be identified it may be wise for prediction to select a range of models across conceptual space in order to reduce the chance of ignoring a plausible model that gives a 'bad' or 'worst case' scenario.

In the next chapter the methods for making this comparison are developed. Thereafter each model type is compared, suggesting ways to diagnose processes and enhance experiments and also indicating where data does not differentiate between processes. Figure 3.7 shows how the models developed or analysed in this thesis populate the framework (each is described in Appendix 1) and Table 3.1 shows which chapters different models appear in.

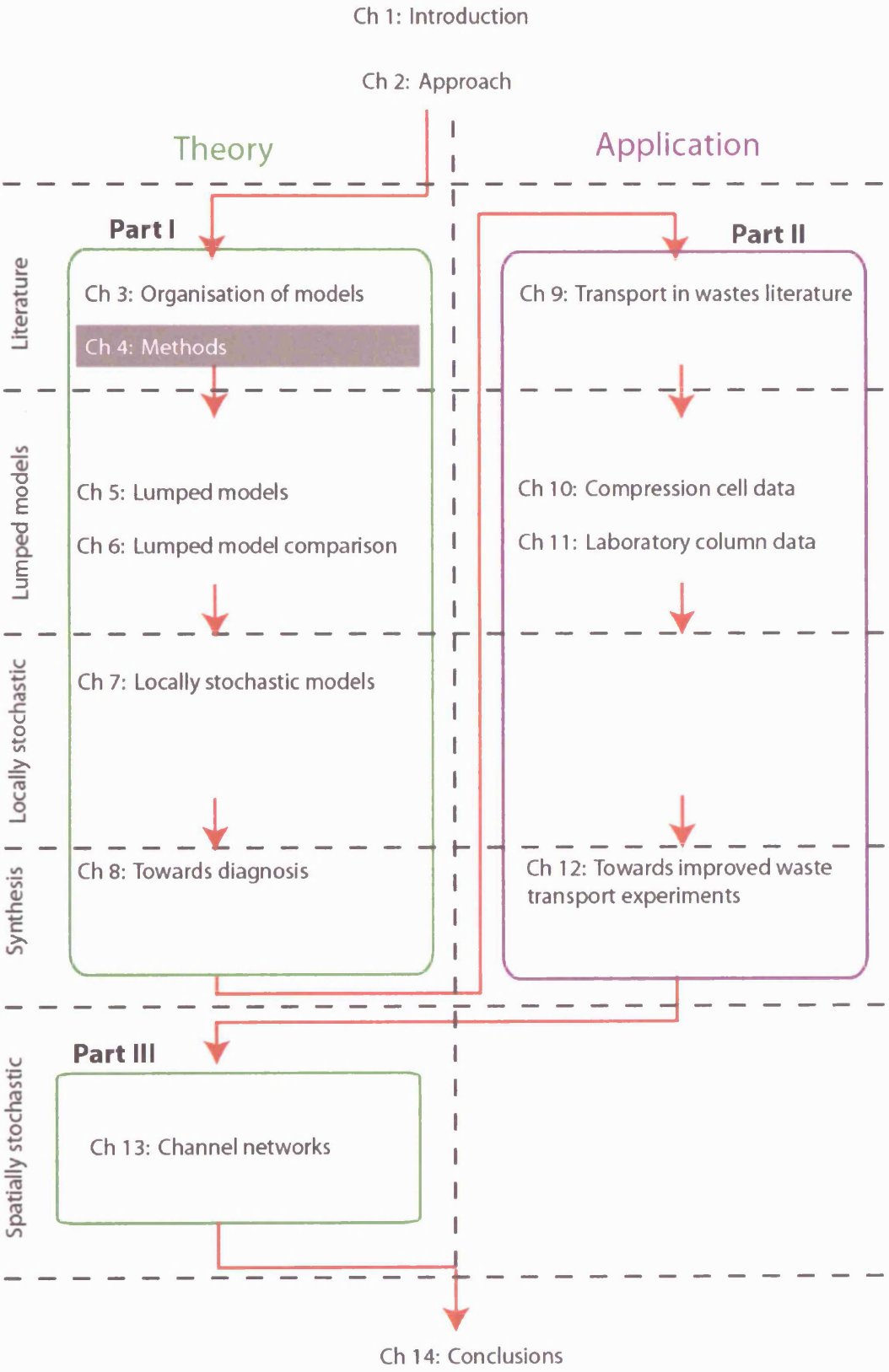


CODE	Description
ARDVARCS	AD model - semi-analytical (piecewise constant flow and inlet concentration)
DPD-PULSE	AD-DP model - semi-analytical solution
HyperConv	Stochastic channel-network model.
LT-INVERSION CODES (APPENDIX 1 and DP1D)	AD-DP model - numerical inversion of 1D transport equations using the de Hoog algorithm
HET-DP	DP model - numerical inversion of 1D DP with heteogeneous block sizes
HET-STREAMS	AD model - summation of many parallel streamtubes
SHARP-SORB	Sorption model - semi-analytical solution to advection with non-linear sorption
SORBER	AD-Sorption model - numerical solution to the ADE with non-linear sorption

Figure 3.7: Primary models populating the framework in this thesis (see Appendix 1 for more details).

Table 3.1: Process categories and their treatment in the thesis.

Category	Process	Details	Chapter
Lumped	AD	Assuming $D = \alpha V$	5
	DP	Different BGFs tested	5
	Sorption	Linear and Freundlich isotherms (not Langmuir)	5
Locally stochastic	Multi-stream-tube	Distribution of tubes only (not bimodal DK)	7
	Multiple-dispersion	Not simulated	-
	Distribution of immobile zones	Various distributions	7
	Distribution of Langmuir isotherms	Discussed (not simulated)	7
Spatially stochastic	Stochastic pathways	Channel network flow model	13
	Stochastic dispersion	Not simulated	
	Stochastic immobile zones	Not simulated	
	Stochastic sorption	Not simulated	


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

4 Model and parameter evaluation methods

To fulfil Objective 1 (*to appraise, assess and organise appropriate models from a pool of possible forms*), Chapter 3 established that there is a need to systematically compare process models in order to be able to identify 'diagnostic' behaviours. This will enable enhanced experimental design in order to better discriminate between models and will also help to identify datasets which do not discriminate. This chapter introduces and justifies a particular choice of measures for comparison.

In common with the choice of models, the choice of measures is generally subjective. The artistry of choosing different measures at different times can appear either arbitrary, or deliberately intended to make a particular point. It is not necessarily practical for most studies to work with multiple measures, but there is space here to do this. It is desirable to know to what extent relationships established with one measure might extend to other measures. This thesis explores the simultaneous use of methods rather than using different methods in series (to do different tasks). The arbitrariness of selection is not completely removed but the methods are at least consistently applied.

There are many possible measures of model performance. From these, three simple and commonly used measures are selected and described. Each of these measures will be systematically applied to the major transport models considered in this thesis (AD, DP and Sorption models) to allow a systematic comparison of models as well as a confrontation of each against waste data. The rationale behind the selection is now explained in Section 4.1.

The primary contribution of this chapter is a simple hierarchy which can be used in analysis of BTC data. The hierarchy is used in analysing the data in chapters 11 and 12. The remaining work in this chapter is a critical review of BTC analysis methodologies. Equation 4.11 was derived here.

4.1 Characteristics of BTCs and flushing data

BTCs and flushing curves are 1D datasets revealing the distribution of travel times of solute passing through the system at one point. This may suggest that parameterisation in terms of times may be desirable.

From this distribution of times there are a few observations that can sometimes be made immediately, for example the first observable arrival and time of peak or peaks. Less obviously it may be possible to detect inflections of slope in the arrival (Figure 4.1).

With a little calculation, simple statistics can be estimated, such as the total recovered mass, mean, spread and skew. The zeroth moment of the BTC gives the mass. The first moment gives the mean travel time. Centred moments are required for obtaining variance, skew and kurtosis. The noise in the system can be examined most clearly for long top-hat inputs when concentration , C, tends to the inlet concentration, C_0, and when $C \approx 0$ long into the tail.

Basic observations (e.g. first arrival time) and statistics are as factually valid as the data they are based on, requiring no additional assumptions. Application of process models requires additional assumptions, both in terms of the models and in terms of how they are applied to the data.

A qualitative appraisal of the shape of the BTC may suggest from experience which process model or models may be most appropriate. This is essentially a judgement-based filter before attempts are made at calibration. Process models will also make certain predictions as to the key observations and statistics. The gradient of the late-time part of the curve (the 'tail') may also be a single 'value'. The residuals achieved between the best-fit process model and the data also give a measure of the system noise, although it is difficult to decouple the model structural error from random measurement errors.

Ward et al. (1998) define a seven-level 'hierarchy of tracer-test analysis'. A rationalised, shorter hierarchy is suggested for this body of work (Table 4.1).

Table 4.1: BTC analysis hierarchy (this thesis).

Level	Description	Comments
1	Transport-connection	Whether any noticeable transport occurs between two points
2	Key features	Arrival time, rising limb inflection time, peak or peaks, truncation time, late time gradient
3	Statistics	Mass recovered, mean, spread, skew etc, noise.
4	Fitting	Fitting of hypothesised process models to the data

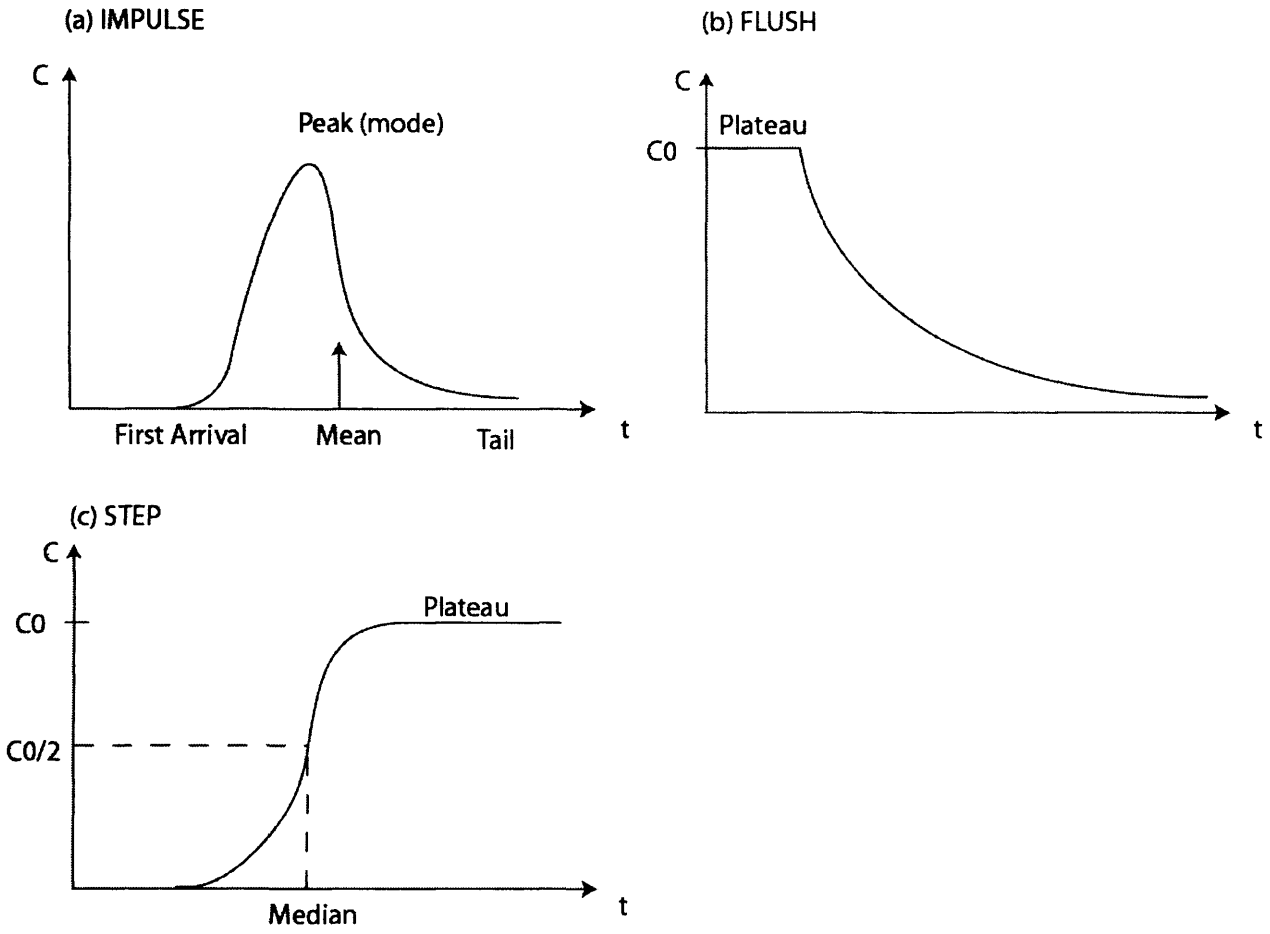


Figure 4.1: Key features of simple concentration data (a) Impulse, (b) Flush and (c) Step responses.

Johnston et al. (2003) and Johnston et al. (2005) produced a taxonomy of BTCs, which were used to classify transport through a wide range of simulated fracture distributions. Monomodal BTCs were a sub-set of a wider range of possible distributions. However, they are important in their own right, since are very commonly experienced with soils, groundwaters and other materials. Unlike more complex BTC shapes, smooth monomodal BTCs have a shape that is relatively simple to match (several empirical representations are able to do this). So a monomodal BTC is both common and relatively poor for use in discriminating models. It is therefore worth more attention and is focussed on here.

Level 1 may not be as trivial as it first appears. In Karstic systems, for example, demonstration of connection between two points may require considerable attention. In general in highly heterogeneous materials Level 1 is not guaranteed. Time-varying connectivity occurs in gassing wastes and highly localised connectivity occurs in channel network systems.

4.2 Level 2: Key features of concentration data

4.2.1 Early time

The ability to predict early time behaviour, first arrival, rising limb and the peak for a tracer test is of relatively low interest here because it is the late-time concentration of flushing which is of primary concern. For this application there remains a question as to how much importance to give the earlier time data, which might reveal or bound aspects of the system. Use of a single value such as the first arrival or peak time to infer process or parameters would imply an enormous weighting for a small number of datapoints. This section reviews to what extent information taken from the early time and peak can inform predictions of the tail.

There are two main difficulties with applying importance to early time data from experiments. Firstly, early time data can suffer from high noise due to hydraulic or other transients caused by starting an experiment (e.g. in the Pitsea compression cell, Chapter 10). Secondly, it is the time at which concentrations are most strongly controlled by the inlet boundary conditions.

'Arrival time' requires an operational definition such as the first measurable concentration (and therefore depends upon the background concentration and detection limit). An arbitrary definition could be adopted such as the time at which $C = 0.001C_0$, in the instance that C_0 is the injected step concentration. If an assumption is made of no mechanical dispersion, the first arrival would enable an upper bound for mobile porosity.

Simple process models are at their weakest at early time. For the impulse response to the AD model $C > 0$ for all z for $t > 0$, which is of course unphysical. There are several scenarios in which a mathematical model of a heterogeneous medium can be shown to tend asymptotically to the AD equation. However, it may take significant time or distance of travel for this asymptote to be reached. Before this, any fit to the AD model will give 'apparent' parameters that will change with time. Field (1999) therefore finds that the arrival time is rendered meaningless by the effect of dispersion combined with measurement difficulties at very low concentrations.

It is tempting to argue that, for an AD model with diffusive exchange (AD-DP), the early time characterises the mechanical dispersion if there has not been time for any diffusive exchange to act. However, there is no theoretical upper limit to the rate of diffusive mass transfer at the leading edge of the invading solute. This is because there is theoretically an infinite concentration gradient between the solute front and the immobile zone. Even ignoring the reliability of this data it is uncertain whether models could be reliably used to identify parameters or processes at early time. Assuming a model with no mechanical dispersion and a

first-order (MIM) exchange gives an unphysical jump in concentration (at first arrival) due to a finite rate of exchange from the mobile zone (in contrast to the diffusive solution which has unlimited exchange).

Moments tend to be dominated by the tail and late-time gradients implicitly concern themselves with the tail. However, the least-squares method is potentially strongly affected by the early time. It would be possible to give these points very low weight (say in proportion to the SD of the data at this point). This latter approach will be adopted as well as simply adopting unweighted least-squares.

In summary, there are good reasons to suggest that early time data will be subject to error due to incorrect assumptions about the boundary-conditions, model error and unreliability due to unsteadiness in the system. Over use or special use of the early time is therefore probably best avoided.

4.2.2 Late time

The concentration at late-time is very important for contaminant flushing applications. A major factor in determining how long it takes a concentration to reach a particular concentration is the gradient of the flushing tail (i.e. the late-time gradient). The gradient is also potentially a further measure in its own right which can be used to help diagnose the process. Analagously to pump-test diagnostics, given a sufficient range in concentration data, a log-log or log-lin plot may be more revealing than a lin-lin plot showing a ubiquitous smooth monomodal BTC or flushing curve.

There has been some discussion as to what extent late-time gradients can be 'diagnostic'. For example, Becker and Shapiro (2000) cite Tsang (1995) as finding that -3/2 log-log slopes are a 'telltale' indicator of matrix diffusion and Hadermann and Heer (1996) attributed a -3/2 gradient at the Grimsel test site to matrix diffusion.

Two key methods are used here for examination of the late time gradient. Firstly, an elegant method by Haggerty et al. (2000) provides an analytical solution which finds the late-time gradient when the mechanical dispersive flux is much smaller than the diffusive return of stored solute. This is adapted in Chapter 7 using nomenclature common to the rest of this thesis (see Chapter 5).

Secondly, a simple method is demonstrated here for establishing the log-log and log-lin late time behaviour. This is done by taking the log of the expression for concentration and

differentiating this. This gives $\frac{d}{dt}\ln(C)$, which is the log-lin gradient. Any terms of the form $\frac{A}{t}$ will be log-log linear with gradient A.

4.3 Level 3: Method of moments

4.3.1 About moments

'Temporal moments' are simple statistics of the BTC. The first four moments have straightforward meanings. Later moments are less clearly visualized and, because of the increasing effect of error, are not used for many practical purposes and are ignored here.

The Nth raw moment is defined as $\mu'_N = \int_0^\infty t^N f(t)dt$ (4.4)

The Nth centred moment is defined as $\mu_N = \int_0^\infty (t-\mu'_1)^N f(t)dt$ (4.5)

Moments are commonly used measures of a curve's shape. These are simple functions that capture the key aspects of shape using only integer power-weighted distances from the mean. It is worth noting that other different measures of shape exist (e.g. spread measured by the average absolute deviation).

The meaning of the first few moments is as follows:

- The first raw moment ('first moment') is simply the mean.
- The second centred moment ('second moment') is the spread (same as the variance)
- The third centred moment ('third moment') is the definition of skew. This measures asymmetry by picking-out differences in values at equal distances either side of the mean.
- The normalised centred fourth moment ('fourth moment') is the definition of kurtosis (it is a measure of 'peakiness').

By taking the Laplace transform of the equation for a moment: $\mu'_N = E\left[t^N\right] = \int_0^\infty t^N f(t)dt$, it is easily shown that the Nth moment can be readily extracted by differentiating the transformed function and setting s to zero. This is often credited to Aris (1975).

$$E\left[t^N\right] = \lim_{s\to 0} \frac{(-1)^N}{\mu'_0} \frac{d^N \bar{f}}{ds^N} \quad (4.6)$$

The centred-moments can then be built up using:

$$\mu_2 = \mu'_2 - \mu'^2_1 \tag{4.7}$$

$$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1 \tag{4.8}$$

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1 \tag{4.9}$$

The method of moments has been used very widely to show that, with careful choice of parameter values, different conceptual models can have the earliest moments matched or made very similar (Bardsley, 2003; Cirpka and Kitanidis, 2000; Goltz and Roberts, 1987; Sanchez-Vila and Carrera 2004 and many others). Note a second use of moments is to 'calibrate' models (Jury and Roth,1990; Woodbury and Rubin, 2000; Cirpka and Kitanidis, 2000).

Moment matching can be deceptive. Ultimately, how well a model manages to describe data is best measured by careful selection of a goodness-of-fit statistic to test the performance of the model in relation to its required purpose. Examples of deceptive behaviours would be the moments for a bimodal arrival with two equally high sharp peaks. The mean would be mid-way between them and the spread would be even more misleading. An arrival which appears to be very sharp on a linear scale, but actually has a very long and low concentration tail, might have deceptively high moments despite this not being obvious visually.

Field (1999) described how transport parameters can be obtained from the moments. In order to be able to do this successfully there needs to be near 100% recovery, which is very rarely the case (BTCs must at some point be truncated), but it may be approximated (for example by assuming the behaviour of the BTC to infinity). Maloszewski and Zuber (1992) used this method to find the arrival time and dispersivity, α, assuming the AD model. Note that this temporal definition of dispersivity differs to the spatial definition given by Gelhar (1993). Spatial moments (Goltz and Roberts, 1987) require spatial data.

The most powerful use for temporal moments is in the comparison of the predicted statistics that would accompany different process models. This method enables discovery, for any given process pair, of which moments can match and what parameter choices are required to do this.

4.4 Level 4: Fitting

There is a lot of choice when it comes to decisions over method of fitting. The methods which are selected for use here are now identified. There is no consensus on what constitutes the 'correct' methodology. It is likely that the 'best' methodology will depend on the context of model use. Ultimately what constitutes a good fit does not boil down to the statistics of a particular fit, but to whether the model can adequately perform what is required of it - in this

case the ability to make a reliable prediction of flushing. A goodness-of-fit measure is a tool towards this, but should not necessarily be made a goal in its own right. The key decisions for fitting are summarised in Table 4.2 and form the basis for the upcoming sections.

Table 4.2: Key decisions for fitting.

Section	Decision	Comments
4.4.1	Choice of what to compare	The primary comparison is with real data, but this may be relatively limited. There are insights to be gained by comparing models to synthetic data.
4.4.2	Choice of objective function	This depends on what is known about the system, what the desirable feature is to match and the convenience and familiarity of the measure.
4.4.3	Choice of weighting	This is potentially somewhat subjective, but weighting to suit a key modelling desire or weighting proportional to reliability of different pieces of data are both rational choices.
4.4.4	Choice of minimisation system	This may be conditioned by the nature of the data (whether multiple minima are expected), but also on what methods it is practical to create or access.
4.4.3.4	Choice of data frequency	This problem is coupled to that of weighting. Sometimes practical constraints mean that there is simply no choice.
4.4.3.4	Parameterisation	Parameterisation affects the weighting.

4.4.1 Choice of comparison

Fitting models against data is the ultimate test of a model (e.g. Figure 4.2). However, for many data sets this will likely be for only one set of conditions, so the effect of the ambient conditions on parameter uncertainty and model identifiability cannot be readily assessed.

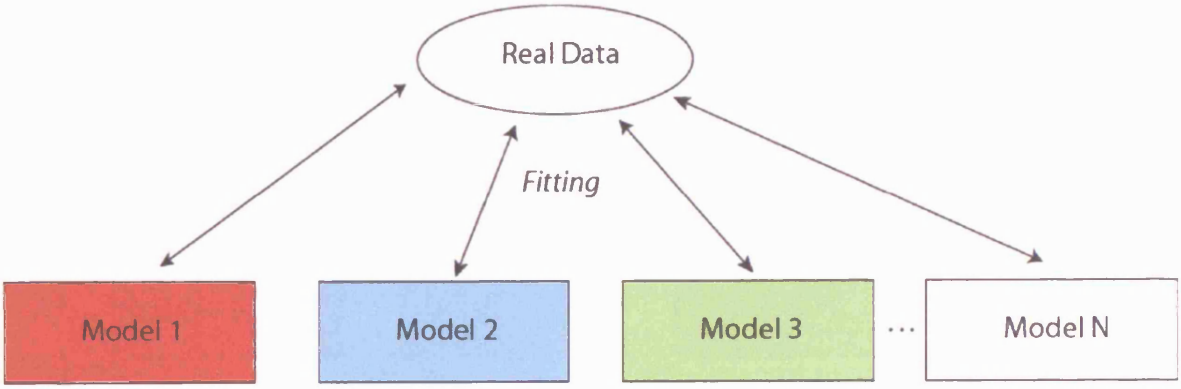


Figure 4.2: Models meet real data.

It is therefore worth considering a few approaches that require only process models and error models. Fitting a model to synthetic data generated by the same model allows an assessment of how accurately the parameters can be estimated even if the model was completely correct (Approach 1 shown in Figure 4.3). This allows assessment of issues such as the importance of

reducing system noise, changing the data frequency and other key experimental operating parameters which may or may not be adjustable. This is addressed in Chapter 5, which examines each process individually.

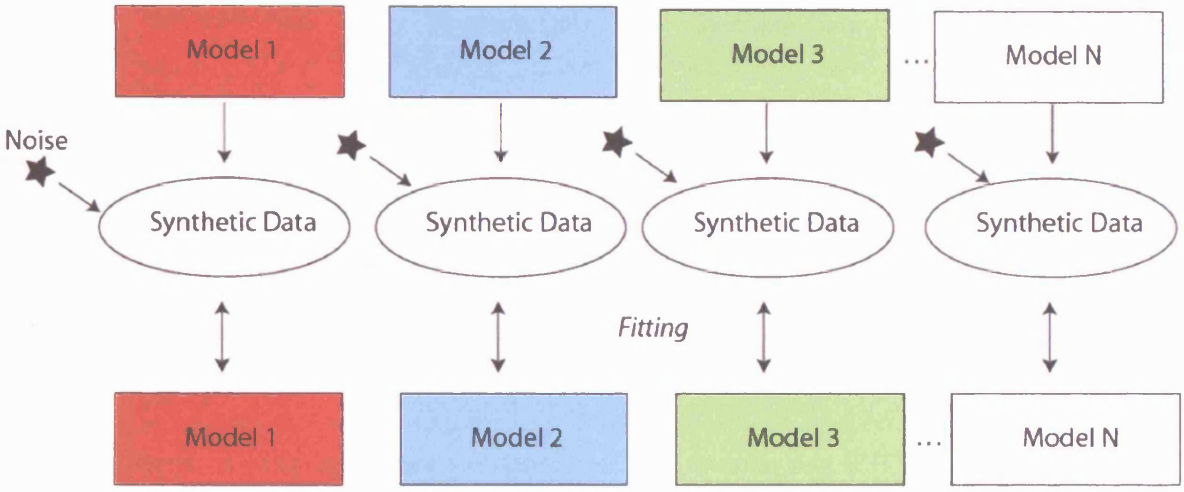


Figure 4.3: Approach 1: Models meet synthetic data (created by the fitting model with noise added).

Fitting a simple model to synthetic data generated by another simple model allows an assessment of the circumstances in which the two models (processes) are potentially indistinguishable (Approach 2 is shown in Figure 4.4). This can be done with or without artificial noise.

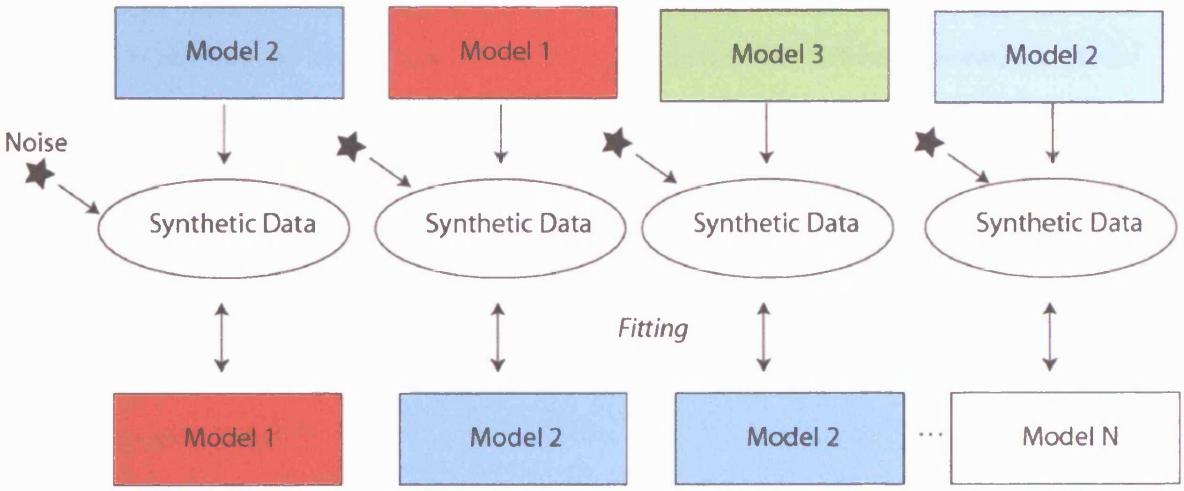


Figure 4.4: Approach 2: Models meet synthetic data (created by a similarly simple model as being used to fit).

Assuming a higher-order model and fitting with a simpler model potentially tests the validity of simpler representations, though this suffers from the difficulty of deciding on what the higher order model should be (Approach 3 is shown in Figure 4.5).

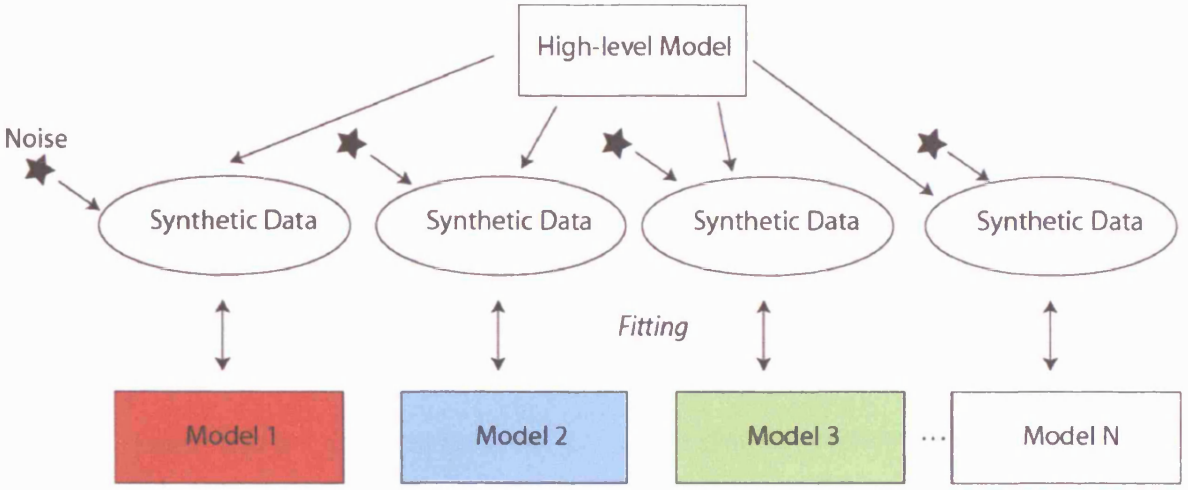


Figure 4.5: Approach 3: Models meet synthetic data (created by a higher order model than being used to fit)

All these forms of comparisons are well-used concepts in analysing BTCs. Bajracharya and Barry (1997) compared fits of the AD model and the ADE plus a first-order (MIM) model to 'data' from a higher order numerical model (i.e. approach 3). Young and Ball (1995) used synthetic data generated by a DP model (with spherical geometry) and fitted with a MIM model (i.e. Approach 2). Adar et al. (1988) used synthetic data with Gaussian noise to test estimation of spatial recharge. Javaux and Vanclooster (2003) added Gaussian noise to a simple AD and fitted with a generalised transfer function. Guswa and Freyberg (2002) took a slightly different approach and simulated a spatially stochastic system with low-K lenses set in a otherwise homogeneous background and matched the first two moments with an AD model and AD-MIM model. A measure of the goodness-of-fit was compared for different ensembles of the simulated system under different generation rules for the stochastic system.

The great advantage of synthetic data is that the system and its errors can be kept simple and controlled. The key issues of similarity and difference under different process conditions can be examined without the expense of repeated experiments.

4.4.2 Goodness-of-fit

A goodness-of-fit quantifies the 'closeness' between a model and a measured BTC in a single number. The benchmark measure is the correlation coefficient, r^2, which is frequently used for comparing BTC data and models.[27] r^2 is supposed to give a guideline for the quality of the fit, for example, "moderate correlation: 0.4 to 0.7, strong correlation: 0.7 to 0.9, very strong

[27] It is built in as the standard in several major soil and groundwater packages; for example, PEST gives a weighted r^2 and CXTFIT gives (confusingly) what is denoted r^2 but is actually Nash-Sutcliffe's efficiency measure, E_N.

correlation: 0.9 to 1.0" etc. The simplest way to express r^2 is as $r^2 = SSR / SS_{yy}$, where $SSR = \sum_{i=1}^{N}\left(E_i - \overline{O}\right)^2$ and $SS_{yy} = \sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2$. In other words it is the proportion of the variance that is explained by the regression:

$$r^2 = \frac{\sum_{i=1}^{N}\left(E_i - \overline{O}\right)^2}{\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2} \tag{4.10}$$

where E_i is each 'correct' modelled value and O_i is each synthetic 'observed' value and $\overline{O}$ is the mean of the observed data. Consider the instance that the model is correct and the only difference is normally-distributed noise, such that $O_i = E_i + e_i$, where e_i is randomly selected from $\mathrm{N}\left(0, \sigma_N^2\right)$. A plot of E_i versus O_i will have a linear regression. The denominator can be re-expressed as:

$\sum_{i=0}^{N}\left(E_i - \overline{O}\right)^2 + \sum_{i=0}^{N}\left(O_i - E_i\right)^2 + 2\sum_{i=0}^{N}\left(E_i - \overline{O}\right)\left(O_i - E_i\right)$. Since the second term is N times the variance and the third term is zero:

$$r^2 = \frac{\sum_{i=0}^{N}\left(E_i - \overline{O}\right)^2}{\sum_{i=0}^{N}\left(E_i - \overline{O}\right)^2 + N\sigma_N^2} \tag{4.11}$$

This is the maximum that r^2 could be expected to reach given a level of noise.

A second difficulty with using r^2 for BTC data is that substantially differing BTCs may have high r^2 values.[28] Therefore using the guidelines for r^2 as a measure of association of a model and data must be used very cautiously.

Due to the squared term, r^2 is most sensitive to outliers (most likely occurring around the peak or around sudden disturbances to the flow). Potentially more problematical for BTCs is that r^2 is insensitive to any additional or proportional differences between the model and data. So, the model could be biased with respect to the data, or the gradient of the regression could be significantly different to 1 without effect.

[28] For example, comparison for an AD model generated with $V = 1$, $D = 1$ and $z = 1$ with an AD model generated with $V = 1$, $D = 2$ and $z = 1$ both with a top-hat tracer input pattern over a time of 5 and points taken every $t = 0.1$ from $t = 0$ to 15 gives r^2=0.97.

What is really needed is a measure of to what extent a model and data are the same. The Nash-Sutcliffe coefficient of efficiency, E_N, is an improvement in this respect and is simply defined as $E_N = 1 - SSE / SS_{yy} = 1 - MSE / \sigma_O^2$, where $\sigma_O^2 = SS_{yy} / N$.[29] This measure is common in the hydrology literature and would be a better 'standard' for comparing BTCs and data than r^2. E_N varies from 1 (perfect fit) to 0 (MSE equal to the variability in the data) to $-\infty$ (poor fit).

A wide range of further goodness-of-fit measures exist. All tend to share the common feature of having a measure of 'distance' which is normalised in some way, giving them a 'portability' between model fits and between datasets. Legates and McCabe (1999) introduced the Index of Agreement which is advantageous in being less sensitive to outliers because it is based on absolute differences. Krause et al. (2005) provided a comparison of several measures. Beven and Binley (1992) described a range of 'likelihood functions'. This does not mean in the strictest sense the same as Maximum Likelihood Theory (but could include it). Gupta et al. (1998) listed several measures which particularly pertain to data where the timing of peak values matter and discussed how multiple measures might be combined in order to achieve several constraints at once.[30]

In the end, the measure chosen is a tool for a specific job and what constitutes the 'best' approach is related to whether the tool helps allow the model to do what is required of it. The primary purpose of the above review is to sound a note of caution over using r^2.

The next sections address the fitting process, which does not necessarily optimise the goodness-of-fit. The least squares methods have advantages; classical statistical methods are available to process them and minimising least squares gives the 'Best Linear Unbiassed Estimator', although the invention of the Bootstrap now rather erodes this advantage (Efron and Tibshirani, 1993). Here focus is on standard methods, so fitting is by minimising an objective (goal) function based on the MSE. Optimising the MSE relates directly to E_N and less directly to r^2.

---

[29] Weglarczyk (1998) gives a useful decomposition: $MSE = \sigma_O^2 \left[\left(1 - r^2\right) + \left(\frac{\sigma_E}{\sigma_O^2} - r \right)^2 + \frac{B^2}{\sigma_O^2} \right]$, where $\frac{\sigma_E}{\sigma_O^2} - r = \sigma_E \sigma_O \mathrm{Cov}(E - O, E)$ and B is the bias. Only in the instance that the last two terms are zero, will $r^2 = E_N$.

[30] This becomes an issue of weighting, which is dealt with in Section 4.4.3.

4.4.3 Weighting and error models

There is little theoretical guidance as to the 'correct' way to weight BTC data. Use of simultaneous measures would be a potentially difficult problem, as the question of how to combine measures is not straightforward. Abbaspour et al. (1997) noted that the determination of weighting factors for adding different objectives affects the results, yet is somewhat subjective. Abbaspour et al. (1997) suggested a multiplicative form of combination in order to eliminate the need for stating weighting factors. However, it is not obvious how doing this is any better than ignoring weights.

Weightings can be used to crudely filter data-points (e.g. removing outliers) or to promote or demote certain regions. For example the Chi-squared measure $\sum_{i=1}^{N}\left(\frac{O-E}{E}\right)_i^2$ is good for reducing the importance of the peaks and increasing the importance of the early time and the late-time tail. This would be a good choice if it happens that the error is proportional to the absolute value of the concentration. In this case this measure also effectively weights each data point in proportion to its error.

This notion applies more generally as a way to relate the weight of a given measurement to the reliability of that measurement. This might require repeated measurement at that particular point, something which is often (but not always) impossible.[31] Alternatively there may be knowledge of the process nature of how error is entering in the system, or good measures of error at the extremes in values, between which an error model might be interpreted.

It is therefore worth considering how error in the BTC may have been generated. Both the real and the modelled systems can be assumed to be made up of different additive (or perhaps multiplicative) error contributions.

Table 4.3: Some possible error sources.

Error	Comment
Conceptual/ Structural	The model may be an inaccurate description of reality. This is very likely and the extent to which this is true is what is of interest, but is obscured by the other errors.
Measurement	There is a good chance that this is random and methods exist for quantifying it, but there is a danger of systematic bias.
Numerical	Some form of calculation (perhaps propagational) error may be in the model Analytical models can be readily shown to hold a given level of accuracy over the entire range of operation, whereas the level of error within numerical models is more difficult to control.
Boundary condition	Error may be due to the true boundary conditions not matching those assumed or measured. This may enter the experiment through poor design of test or perhaps

[31] Aside from repeating the experiment, measurement of multiple outlet points is one way (although these are not repeats), another way is split samples (this is potentially useful for understanding the sampling and analytical contributions to error).

Error	Comment
	poor control of the injection rate. Mass balance is the first most obvious test of this.
Initial condition	Error may be due to the true initial conditions not matching those assumed or measured. This will be introduced if the initial conditions are not what they have been assumed to be.
Process conditions	The system may not be constant with time. Particularly difficult to control is flow rate, but other variables such as porosity may vary with time too.

A simple error model would lump all the contributing errors and imagine that an observed data point comprises the true value added to an error (i.e. $O_i = E_i + e_i$). To assume that this error is completely random requires an assumption that the non-random contributions discussed above are very small in comparison to the random ones. The error, e, could potentially be described by a statistical model such as the Gaussian model, where e is randomly selected from $\mathrm{N}\left(B,\sigma_N^2\right)$ where B is zero unless there is a bias. By assuming that the noise is the same for all datapoints, an estimate of noise is obtained by using $\sigma_N^2 \approx \sum_{i=1}^{N}\left(O-E\right)_i^2/(N-M)$, where N is the number of data points and M is the number of parameters (van Genuchten, 1999). [32]

Establishing the magnitude of the error at a particular point is more difficult (σ_i^2 for each point could be achieved by repeat samples). In this instance, a weighted SSE can be expressed as $\chi^2 = \sum_{i=1}^{N} w_i\left(O_i - E(x_i;a)\right)^2$, where weightings are set such that $w_i = 1/\sigma_i^2$, otherwise, the error could simply be assumed. More complex error models might be imagined (e.g. auto-correlated errors), but are not considered here. Hereafter, the observed concentrations are written C_i and the modelled ('expected') concentrations are written $C(t_i;a)$, where comparisons are taken at times t_i and for a vector of parameters, a.

4.4.3.1 Unweighted least squares

Assuming noise is equally distributed across data points as a random error, e, then $C = C_B + e$, where C_B is concentration before noise added and e is drawn from a single distribution, $\mathrm{N}\left(0,\sigma_N^2\right)$. The magnitude of every $C_i - C(t_i;a)$ has the same pdf. Giving each comparison point equal weight (i.e. setting all $w_i = w$) is commensurate with each point having equal uncertainty and the weighting no longer needs to appear in the sum. This gives:

[32] Note that $(N - M)$ is the number of degrees of freedom.

$$\chi^2 = w\sum_{i=1}^{N}\left(C_i - C(t_i;a)\right)^2 \tag{4.12}$$

Substituting a normal error model for $C_i = C_B + e$ gives: $\chi^2 = \sum we^2$. The choice of w is 1 means that all data points have the same expected contribution to χ^2. The drawback of assuming $C = C_B + e$ is that for $C_B \approx 0$, unphysical negative Cs may be predicted. Equally for a long step input, once $C_B \approx C_0$, predictions of unphysical $C > C_0$ can occur.

4.4.3.2 Unweighted least squares (bootstrap)

Instead of assuming Gaussian error, any number of different distributions could be used and applied independently to each point, assuming again that the error is independent and identically distributed (iid) at any point. It is desirable to have a good justification for the underlying stochastic process. For example, both additive (normal) and multiplicative (lognormal) processes arise from the Central Limit Theorem (CLT). If such a justification is difficult to make it may be preferable to infer a distribution from the data and use the bootstrap methodology (Efron and Tibshirani, 1993; Press et al., 1992). The bootstrap can be coded by storing all the errors between the data and the best-fit in an array and randomly sampling from this array. None of the errors are removed from the array, so each could potentially be chosen more than once (typically $1/e \approx 37\%$ points are twice sampled). This strategy is known as 'sampling with replacement'.

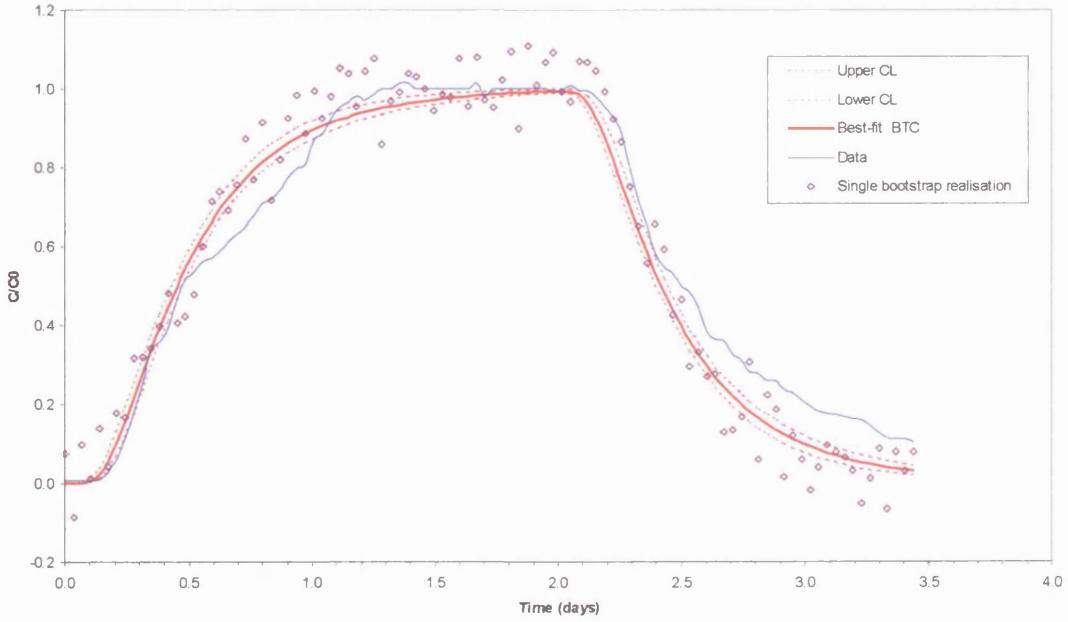


Figure 4.6: Bootstrap realisation based on the difference between the best-fit and a BTC for bromide passing through a column of Mercia Mudstone clay.[33] The best-fit line and the 95% Confidence Levels (CLs) either side are shown. Notice that the data are not always within the 95% CI.

[33] Data provided by Dr Anne Stringfellow from the University of Southampton.

Despite the large simulated variance that is achieved by the example Bootstrap in Figure 4.6, it is clear that substantial portions of the BTC lie outside the 95% CI. The assumption of uniform distribution of error gives negative and overly high estimates of the plateau. Other errors, perhaps due to initial or boundary conditions and very probably model structural error, are to be suspected.

4.4.3.3 Lognormal weighting

Assuming that noise is generated by $C = C_B.10^e$, where C_B=concentration before noise added, e is randomly generated from a Gaussian, $\mathrm{N}\left(0,\sigma_N^2\right)$. This is a **relative** error which scales depending on C_B. Taking logs gives $\log C = \log C_B + e$. So when viewed on a $\log C$ axis the error will appear to have the same variance at all concentrations (analogous to the unweighted case). Note that this method avoids the predictions of negative C that occurs for $C = C_B + e$, but does not avoid the possibility of $\frac{C}{C_B} > 1$. Since $\chi^2 = \sum_{i=1}^{N} w_i C_B^2 \left(10^e - 1\right)^2$, the choice of $w_i = 1/C_B^2$ ensures that all points have the same expected contribution to χ^2.

4.4.3.4 Implicit weighting by sample frequency or parameterisation

Some datasets have a variable frequency of sampling and therefore an implicit weighting. This data can potentially be filtered and interpolated to give a new frequency, although it is often desirable to use all the data.

In terms of designing a new experiment, the frequency can be seen as a choice. Kinniburgh (1986) found that there is no simple answer to the question of possible optimal distribution of data points. He suggested as a reasonable approach the even spacing of data points on a linear scale if the error is independent of concentration, and along a log-scale if the error tends to increase with concentration.

For Part I of this thesis, the sample frequency is held constant. Control of weighting can then be entirely by the weighting function.

Kinniburgh (1986) also showed that minimising the least squares plotted in different ways is similar to applying different weightings. The point pertinent to BTC fitting is that changes of parameterisation (plotting on differently-normalised or dimensionless axes) will potentially

affect the effective weighting. For the theory section the time axis is not normalised, so as to represent the most basic case.

4.4.4 Minimisation and error space

Finding the point of minimum error requires adjustment of the parameters until the particular (weighted) error is minimised. This activity is synonymously referred to as 'fitting', 'calibration', 'parameter identification' or as an 'inverse method'.

Automated calibration of models is now very common. There are various public-domain programmes available such as PEST (Doherty, 1994). Carrera et al. (2005) summarised a range of algorithms for finding the minimum.

Unlike the choice of model, objective function and error model, the choice of minimisation method does not critically affect the results, unless dealing with a 'pathological case'. The choice relates more to numerical efficiency. Therefore the near-standard gradient method of 'modified Levenburg-Marquardt' is selected for use here (both coded into new FORTRAN programmes but also by using PEST). Details of this method can be found in Press et al. (1992) and in the documentation of the IMSL BCLSF subroutine (Visual Numerics, 1997). An advantage of the Levenburg-Marquardt method is that the covariance matrix is calculated as an inherent part of the minimisation (although the BCLSF subroutine rather abnormally does not output this). From this matrix an estimate can be made of the confidence region (CR) around the minimum.

PEST (Doherty, 1994) adjusts the parameters of a model in order to minimise the weighted least squares difference between the model and data. This is done by setting up a 'control' file which informs PEST of the name of the model executable file and the model input file. So-called 'template' files are used to show PEST where the input parameters are within an input file and so-called 'instruction' files are used to tell PEST where the required output from the model can be taken from. Starting from a set of parameter values (which must be specified in the control file), PEST then uses a nonlinear estimation technique to find the values giving the local minimum in the sum-of-squares error. The user has a number of decisions to make:

- Which data to use
- What weightings to give that data
- Which parameters to fix and which to vary
- Whether to vary parameters on a linear or a logarithmic scale
- What range to bracket parameters within
- What initial parameter values

- Whether to 'tweak' any of the parameters relating to the minimisation algorithm based on any prior knowledge of the nature of the response surface.

There are a number of useful statistics generated by PEST once the code has met its convergence criteria:

- The best-fit values of the parameters and their 95% confidence intervals (CI)
- The weighted r^2 of the best-fit point
- The parameter correlation matrix
- The parameter covariance matrix
- The eigenvalues and eigenvectors of the covariance matrix.

It is worth noting that the 95% CI is computed using $\sigma_N^2 \approx \sum_{i=1}^{N}(O-E)_i^2/(N-M)$. It should be emphasised that this is purely an estimate of the underlying noise (which is assumed to be normally-distributed). In the Monte Carlo method in Chapter 5, σ_N^2 is assumed rather than estimated and the noise is normally-distributed.

The biggest weakness for gradient methods, in common with other minimisation methods, is the possibility of numerous minima. A gradient method could become trapped in a local minimum potentially missing a more important 'global' minimum. This is a particular problem for highly parameterised systems. In this instance other methods may be required (potentially in conjunction with gradient methods). One such method is SUFI, which is based on a random search (Abbaspour et al. 1997; Abbaspour et al. 1999; van den Daele 2005). SUFI suffers from fewer convergence problems than gradient methods and can in principle find several local minima and the global minimum.

However, for the situations in Chapter 5 in which synthetic data are used, examination of the error space rarely revealed multiple minima. Conventional gradient methods were therefore employed.

As well as checking on the number of minima, examination of the error space as a surface gives a clearer picture of the effect of different parameters. A two parameter system can be examined on a contour plot, and a three parameter system can be visualised as a series of 'slices' where one parameter is held constant for each slice, or as surfaces of constant goodness-of-fit . Four parameters and beyond are more difficult to visualise. Some schematic 2D plots of error space are shown in Figure 4.7.

The shape of the error space may further reveal:

- How far from the minimum the error contours be regarded as approximately elliptical?
- What (if any) the directions of correlation between parameters are?
- In the instance of more intense correlation, whether there are lines or arcs of non-uniqueness?

It is worth knowing whether the error contours are approximately elliptical. If the contours are not elliptical, the CR predicted from the covariance matrix may be a substantial underestimate of the true uncertainty in parameters.

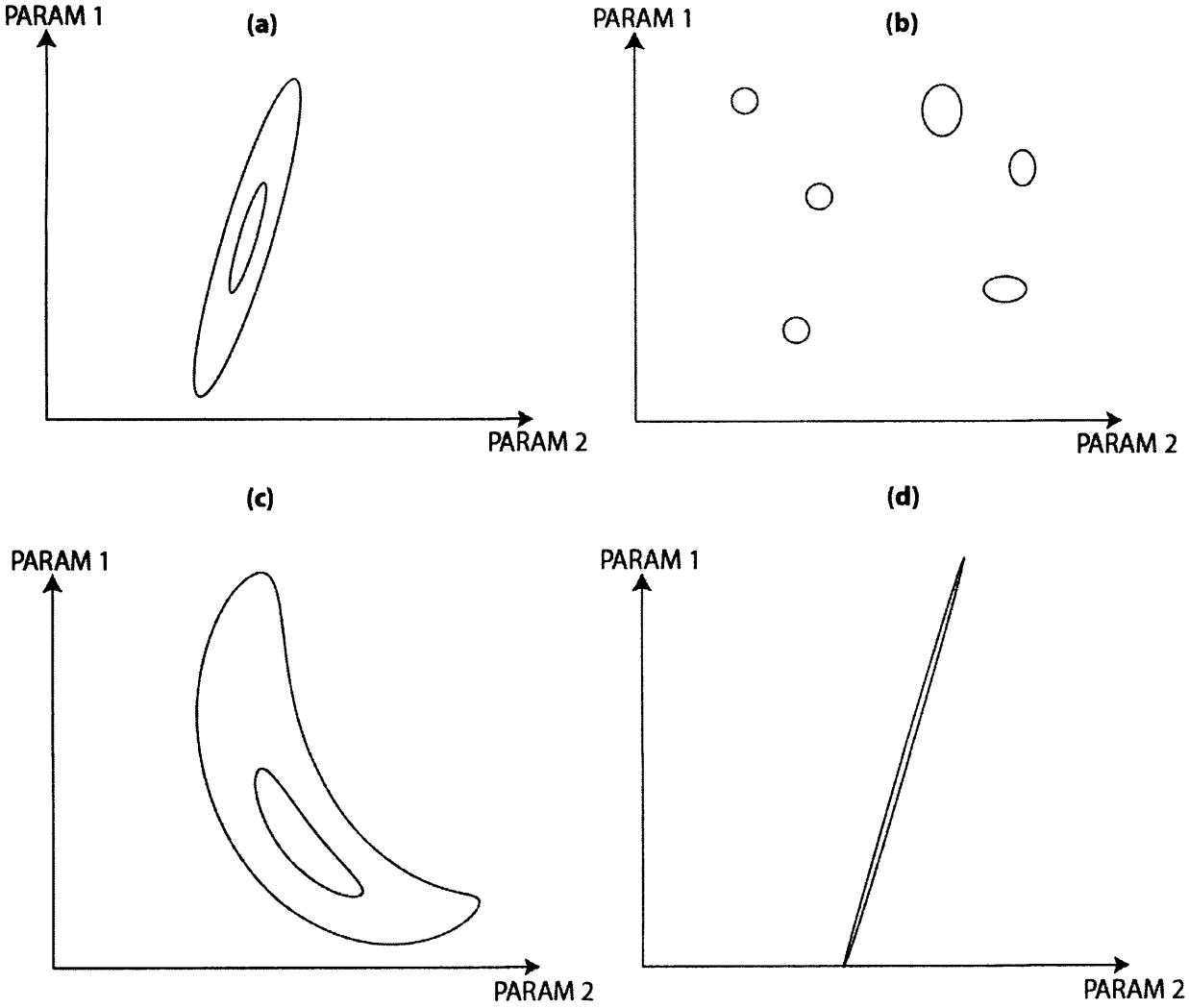


Figure 4.7: Hypothetical error space shapes (arbitrary singe contours) projected into 2D (a) 'well-behaved' elliptical, (b) multiple minima, (c) non-elliptical (curved), (d) line

A FORTRAN code (ERROR-SPAC.F90) has been written here to contour any set of parameters for any given process model and goodness-of-fit measure. It calls the Matlab® engine for production of graphics (see Appendix 1).

The limitation of plotting the error space, aside from the computational burden, is that it is difficult to know which contour of error forms a reasonable boundary for how far estimated

parameters could vary away from the best-fit. Clearly there is little interest in the error-space so far from the minimum where the BTC bears little resemblance to the target. What is required is a method for identifying sensible contours of error. The Monte Carlo (MC) method provides this.

A further advantage of the MC method is that it avoids the use of the covariance matrix for describing the confidence region and then relying on a linearisation which may not necessarily be valid. Standard codes such as PEST are unable to test for this validity, whereas a full MC simulation can provide full verification of the estimated region. PEST is used later in this thesis, but suffers from the further disadvantage relative to the Monte Carlo method that it cannot be adapted to accept different error models.

Chapter 10 describes considerable difficulty in fitting DP models to real data, a problem which was primarily attributed to the minimiser struggling with strong directions of correlation (non-uniqueness). Wright (2002) observed very distinct directions of correlation when fitting DP models to ASR data. Indeed the 'valleys' were so elongate, that it was conjectured that elliptical error space might not be accurately estimated from the covariance matrix. With this experience in hand, difficulty was therefore anticipated in generating reliable results for systematic work with DP models

In summary, the need for prediction, the inherent noisiness in BTC datasets and the strong possibility of lines or directions of strong correction combine to make a case for development of and use of a MC calibration code. This is employed having taken the decisions that are described in this section and that are outlined in Table 4.2.

4.4.5 Monte Carlo methods

4.4.5.1 Monte Carlo code description

A FORTRAN MC simulator "MONTY.F90" has been coded based on the methodology given in Press et al. (1992: Section 15.6). There follows a brief summary of the method, but for more details refer to Appendix 1.

The outline of the MC method is as follows:

1. Read in BTC data or generate with a model.
2. Store the parameters (a_{true}) for which the model was generated or use the Levenburg-Marquardt fitting algorithm to find the best-fit parameters for this BTC.

3. Compute χ^2 at the best-fit point.
4. Generate random normal, log-normal or bootstrap noise and add it to the model (or modelled best-fit) BTC to create a synthetic data-set.
5. Use the Levenburg-Marquardt fitting algorithm to minimise chi-squared (χ^2) of the synthetic data and find the new best-fit parameters.
6. Store the parameters and repeat N times. It is rather arbitrary what N should be selected. 10,000 is adopted here as a number commonly cited to be adequate.
7. Calculate the Hessian matrix invert to get the covariance matrix and then calculate eigenvalues and vectors.
8. Compute the CIs of the different parameters from the covariance matrix.
9. Calculate the locus of the confidence ellipse.
10. Count the number of points within the ellipse and check that the ellipse indeed describes the correct CI.
11. If required, automatically adjust the size of the confidence ellipse until the required number of points lie within it. This is done here by the bisection method (see Press et al. (1992) for details).

The basic MC scheme is encapsulated in Figure 4.8.

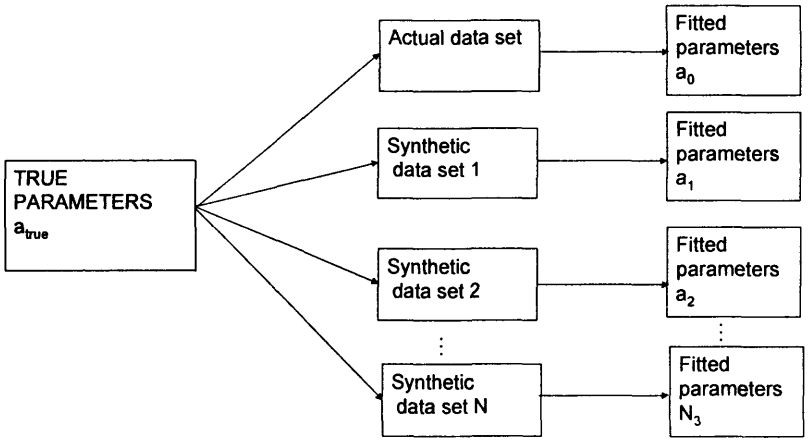


Figure 4.8: Monte Carlo method (based on Press et al., 1992).

MONTY.F90 copes with one, two or three parameters, but could in principle be generalised for larger numbers. Use of two dimensions has the greatest application to the simple process models here.

If the shape of the error ellipse is known to match what is predicted based on the Hessian, the ellipse can be plotted and CIs can be found in a single step without needing thousands of Monte Carlo (MC) realisations.

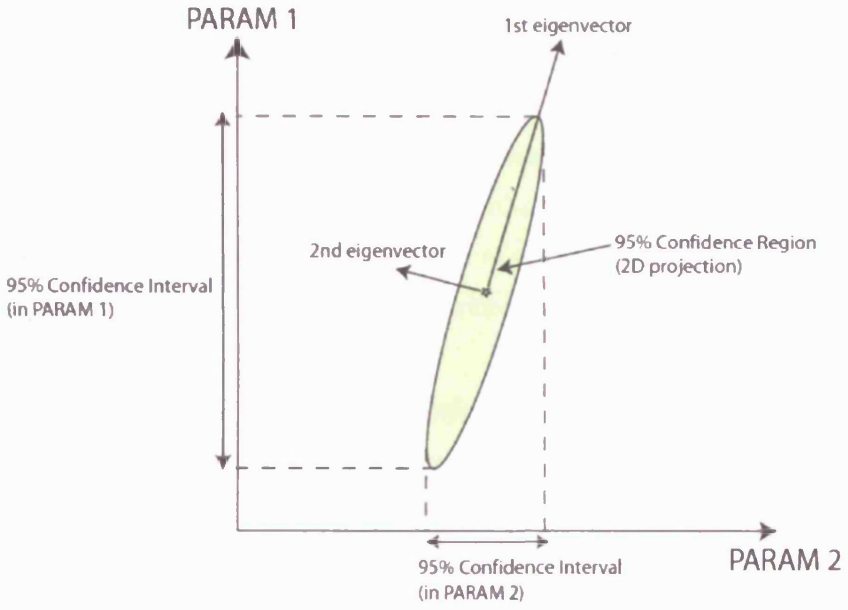


Figure 4.9: CR, CI and eigenvectors of elliptical error space.

4.4.5.2 Obtaining confidence intervals (CIs)

A useful result is for the confidence region (CR) projected into one dimension to give a CI. This is given by Press et al. (1992) as $\delta a_0 = \sqrt{\Delta\chi^2}\sqrt{C_{00}}$, where C_{00} is a term in the covariance matrix, $[C_{ij}]$. As $\Delta\chi^2$ is distributed as χ^2, its value can be established given the number of degrees of freedom and the required confidence level. These are conventionally tabulated for σ=1,2 etc. However any CI could be chosen and can be calculated using the incomplete gamma function (see Appendix 1). Figure 4.9 shows a 2D CR projected into CIs for single parameters as well as the principal directions.

4.4.5.3 Final comments

MC analysis is computationally expensive. The model is called N times multiplied by the number of steps required to find the minimum. A saving can be made by reducing the number of datapoints, M. The default used in this thesis is M=50. For slower numerical models the computational expense can be a real burden, so only fast semi-analytical codes are submitted to MC analysis here.

However, unless the accuracy of the elliptical prediction using the covariance matrix is known, there does not exist any better way of establishing the CR. Because of this, the CR from the covariance matrix becomes only a semi-quantitative guide. This equally applies to use of eigenvectors (the directions of the principle axes of the ellipse) and eigenvalues (the relative length of the semi-axes). The direction of the largest eigenvector will be the direction of greatest parameter (or parameter group) insensitivity. The ratio of the largest to the smallest eigenvector is a measure of the extreme aspect ratio of the ellipse, and can be screened for instances where there is near-complete non-uniqueness (when it gets very high). Use of eigenvalues is particularly helpful in multi-parameter problems where plotting is difficult and where projecting into a smaller number of variables may potentially be misleading. However, they suffer from precisely the same weakness as for the estimated CIs in terms of accuracy.

Log-transforming the variables can help to make a more well-conditioned (i.e. close to circular) error ellipse. When transformed back, this gives a curved CI which is no longer elliptical. Parameter values are also kept above zero.

It would be useful to both highlight where CRs that have been established from the covariance matrix are expected to be reliable and also to establish where parameter uniqueness and identification problems may arise. For each simple process model MC realisations are performed to investigate this in Chapter 5. A complete coverage of all parameter permutations was not possible, but the principle is demonstrated.

A weakness of the MC method as given in Press et al. (1992) is that the goodness-of-fit of the best-fit point for each realisation is not reported. A distribution is given of parameters that achieve the minimum in error, but without indication of what the absolute value of this error is. It is worth evaluating sample BTCs and outputting goodness-of-fit measures at points on the locus of the chosen CR.

A further shortcoming is that the location of the points in the MC scatter is only as good as the fitting routine. Whether the fits are repeatable is readily tested by randomising the starting guess of parameter points.

Methods such as GLUE (Beven, 1993) work without requiring any assumption of error structure or use of minimisation routine and explicitly take into account the goodness-of-fit of each realisation for use in prediction. In one go this method removes the main disadvantages of what is done here. However, GLUE does not share the powerful potential advantage of CIs being available based on the best-fit and requires 'prior' pdfs of parameters to be estimated; something which is notoriously difficult to do.

The Monte Carlo method does not assume that the shape of the error surface is elliptical. This allows the accuracy of the confidence interval estimation using this assumption (via the Hessian) to be established. The reliability of the CIs given by PEST (and any other code based on the same principle) can therefore be tested in this way.

4.5 Units and parameterisation

In order to remain generic all models are presented without units. Any set of internally consistent units could be applied, for example Table 4.4.

Table 4.4: Examples of consistent unit sets.

Variable	Dimension	A few example choices of units		
V, K, q, V_m	[LT^{-1}]	m/d	cm/d	cm/s
D_a	[LT^{-2}]	m/d^2	cm/d^2	cm/s^2
Lengths: L, z, α, a, b	[L]	m	cm	cm
Area: A	[L^2]	m^2	cm^2	cm^2
Volume: Vol	[L^3]	m^3	cm^3	cm^3
Times: t, t_{cb}, t_{cf}, t_a, t_a', t_D, t_M	[T]	d	d	s
MIM coeff κ	[T^{-1}]	d^{-1}	d^{-1}	s^{-1}
Mass, M_0	[M]	mg	g	µg
Concentration: C_0, C_i, C	[ML^{-3}]	mg/m^3	g/cm^3	$\mu g/cm^3$

Thought has been applied to the 'best' way to parameterise different models. There is considerable choice and there are several permutations of different times and dimensionless numbers. There are advantages in normalising in order to achieve dimensionless plots, however, the clarity gained through having unaltered units is preferred here.

4.6 Summary

A revised hierarchy of tracer breakthrough curve analysis was suggested in Table 4.1. Here, one method is selected from each of the top-three levels. This section describes the choice of measure.

Table 4.5: Summary of analysis level and usage in this thesis.

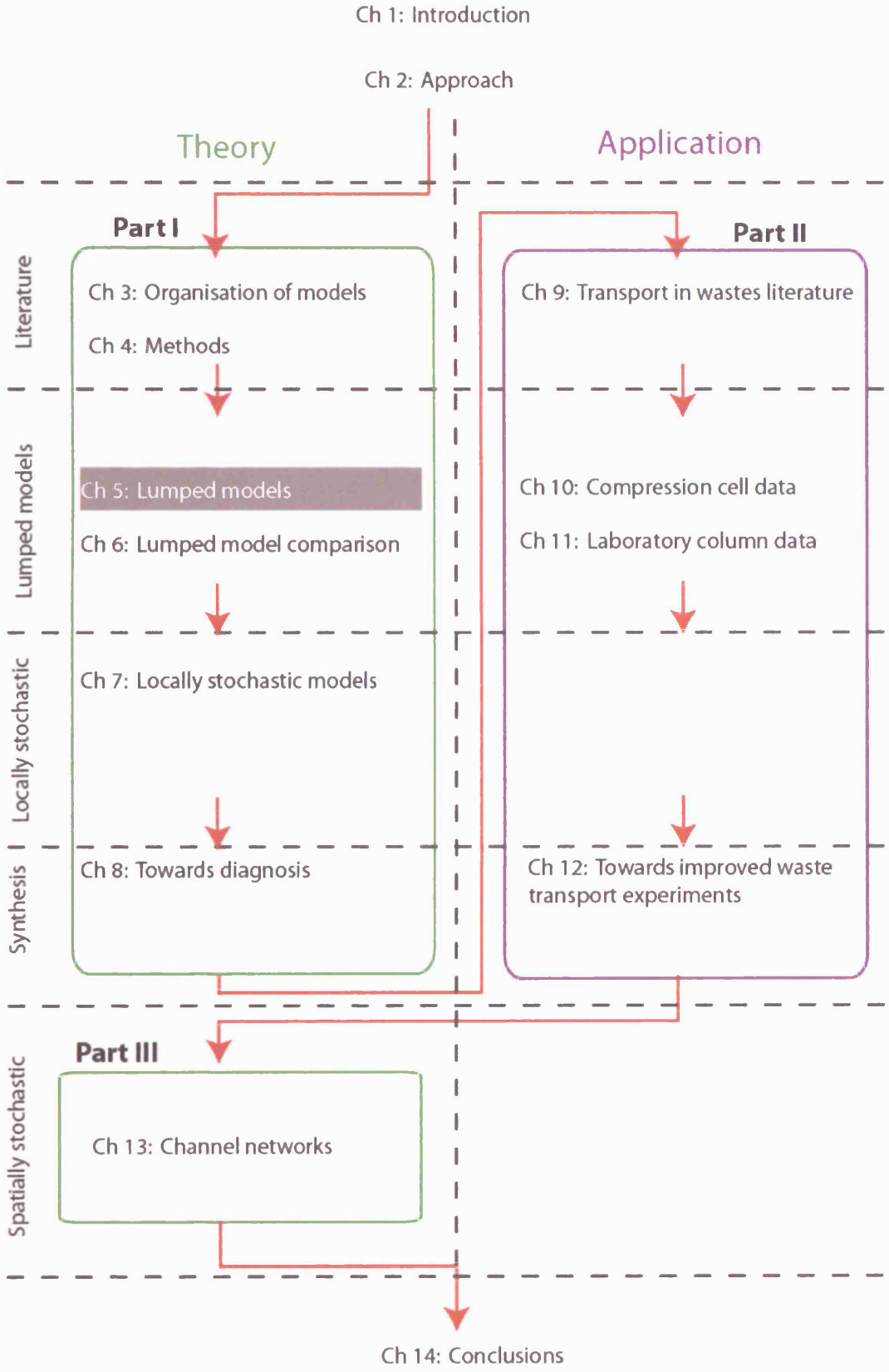
Level	Description	Comments
1	Transport-linked	A prerequisite for this work to occur.
2	Key features	Late time gradient
3	Statistics	Moments (mean, spread etc)
4	Fitting process models	Least-squares and Monte Carlo

None of these methods is without problems. Table 4.5 summaries the advantages and disadvantages of each. The advantage of using multiple methods is that it combines the strengths of all without being too badly affected by any one single weakness. Using methods in parallel avoids using a series of methods to furnish arguments at a particular time.

In conclusion, this chapter has selected and appraised three methods for comparing models - least-squares fitting, method of moments and late-time gradients. This is the first step towards achieving Objective 1: *to appraise, assess and organise appropriate models from a pool of possible forms*. The next step is to provide an organised set of models and then to systematically compare each using these methods (Chapter 5).

Table 4.6: Appraisal of selected methods.

Method	Advantages	Disadvantages
Moments	• Commonly used as a method of model comparison. • Produce analytical results for theoretical process models. • Provide one way to select dominant processes . • The spread can be thought of as a measure of dispersion. However, formally it is the spatial spread that has been given a definitional relationship to 'dispersion' (Gelhar, 1993). • Very easy calculation in Laplace space of what could be quite difficult integrals. Ready-matching of equations that are not readily inverted. • The matching of systems with several parameters can be done this way. It is difficult to express such matching graphically when using least-squares, for example.	• Real data may be truncated, but the moments are defined from zero to infinity. In order to apply this method to real data, the all-important tail has to be guessed. • The 'moment arm' of large t values is potentially large, especially for a highly skewed distribution, so small absolute error can play a large role in the tail. • As moments are the statistics of BTC shape, they can potentially be used as 'weak' criteria for matching. However two BTCs with identical spread can be completely different shapes. It is tempting but dangerous to use this method as a proxy for fitting. • Counterintuitive effects (e.g. dispersion might increase at the same time the peak gets steeper, as illustrated in Chapter 5).
Late-time slopes	• Important for prediction of long-term future. • Single simple number. • Less affected by early time or peak noise and error.	• Only one number • Data not easy to collect
Least squares	• Potential for 'complete' curve comparison.	• Weighting, choice of comparison points and choice of least-squares as a goodness-of-fit measure are all relatively arbitrary. • For number of parameters or independent parameter groups >2 it rapidly becomes difficult to visualize parameter space.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

5 Lumped models

Lumped models at different points in the process triangle (Figure 3.1) are now examined. Figure 5.1 shows how the sections of this chapter relates to the triangle.

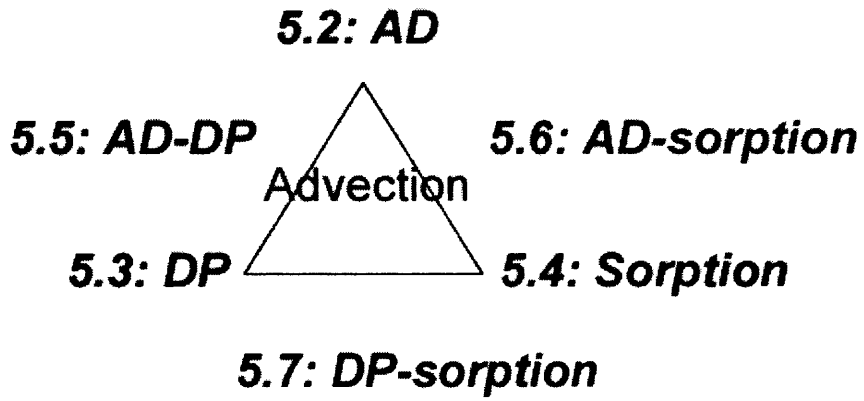


Figure 5.1: Lumped models appearing in Chapter 5 (by section) - combinations of the Advection-Dispersion (AD), Dual Porosity (DP), Advection (A) and sorption processes.

Primary amongst the new contributions in this chapter are:

- Systematic comparison of the lumped process models shown in Figure 5.1 by the simultaneous application of the method of moments, late-time gradient analysis and Monte Carlo analysis of least-squares fitting. No such comprehensive and thorough matrix of comparison has yet been performed, although many of the individual aspects of each model have been dealt with in the literature (but in isolation).
- Examination of the shape of the error space for the AD and DP processes by Monte Carlo analysis
- A unification of approaches and models into a single logical nomenclature based on characteristic timescales and generalised by use of 'Block Geometry Functions'.
- Calculation of late-time log-gradients for the AD, AD-DP and Freundlich sorption model is new.
- Calculation of late-time log-gradients for the DP process is based on the methodology of Haggerty et al. (2000), although it is based on diffusive geometries (rather than the multi-site MIM version of Haggerty et al. (2000)) and now incorporates the generality of the BGF.
- Discovery of the theoretical possibility of characterising the block timescale for DP systems by using short-duration tests
- Examination of the effect of Freundich sorption at low concentration
- Suggestion of a simple logical diagnostic test to reveal the presence of non-linear processes

Review material in this chapter is summarised as:

- Introductory and general observation sections for each model (Sections 5.2.1, 5.2.4.1, 5.31, 5.4.1, 5.4.2, 5.4.3, 5.4.4,)

- The 'suite' of analytical models presented in 5.1.2 are all solutions previously derived (they are re-derived and re-expressed here in order to give clarity, generality and flexibility).
- The moments of all the models given here have been previously derived, although with non-unified parameterisation and without the generality afforded by the use of BGFs in the treatment here.

5.1 *Framework of AD-DP models*

In order to achieve Objective 1 (*to appraise, assess and organise appropriate models from a pool of possible forms)*, an organized set of models in the 'pool' is required.

A set of entrants to the pool is now developed based on the AD and DP models. In Section 5.4 sorption models are added, completing the triangle of process types suggested in Figure 3.1. Later, in Chapter 7, representations of heterogeneity are added to this basic structure, yielding yet more possible models in the 'pool'. Having developed the family of models, the first process analysed in detail is AD, in Section 5.2.

The same basic approach was described by Sardin et al. (1991). They built up complexity from the same equation, generating increasingly complicated and parameterised models. In contrast here several different conceptual models are created at a similar level of abstraction.

5.1.1 Framework

The different processes and process combinations that are examined in this thesis are shown in Figure 3.1. Because the aim of this work is to make a process-based comparison and discussion of diagnosis, the use of numerical codes is minimised and analytical solutions are used wherever possible. This is because, being more complex, numerical codes tend to require a greater level of documentation, error analysis and verification. The coupling of non-linear sorption and DP is not examined.

These processes need to be examined under certain simple conditions:

- Impulse (temporal Dirac delta) input.
- Step input.
- Flushing of solute (initially uniformly distributed within the medium).

Consideration of more general inputs (flow rate and concentration varying with time and non-uniform initial concentration) are considered in Chapter 10 in relation to data, as well as in Chapter 8, in relation to intermittent flow tests.

5.1.2 Laplace transform (LT) solutions

The Laplace Transform (LT) approach used here has the advantage that it provides a family of solutions. This section describes the same system of equations as appeared in Barker (1982) but it is reparameterised using the later 'block geometry function' nomenclature of Barker (1985a&b), thereby describing *any* diffusive geometry. Barker (1982) considered the specific case of a slab geometry. The outlet boundary was assumed to be infinite in Barker (1982), but here finite boundary conditions are also considered.

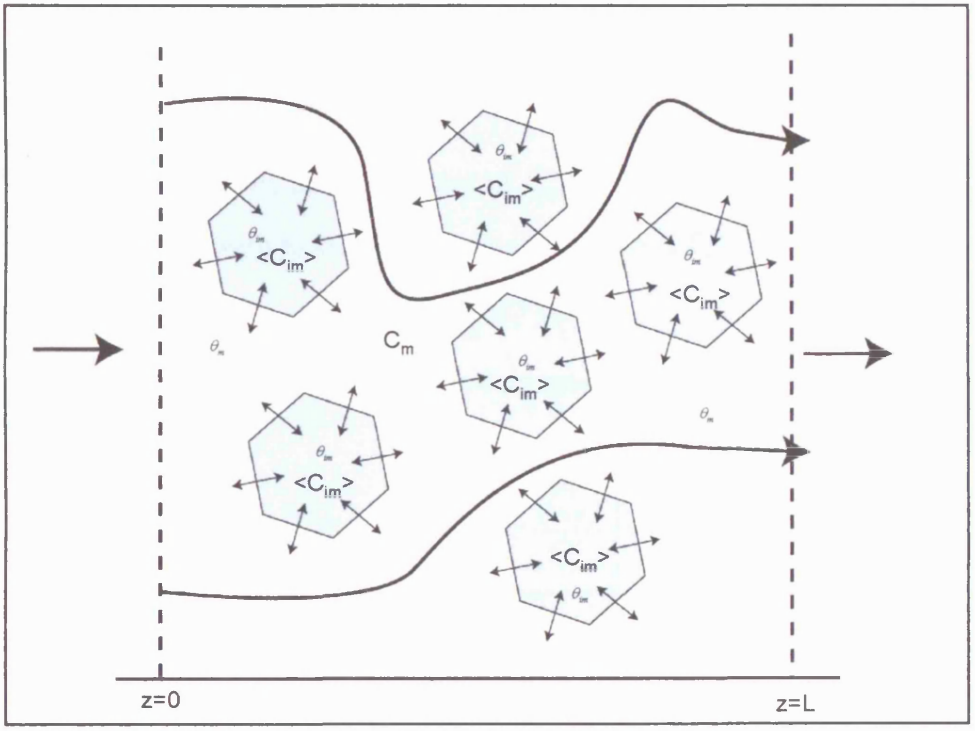


Figure 5.2: Basic transport system assumed in Chapter 5.

The transport system assumed here is shown schematically in Figure 5.2. The following assumptions are made: (i) the impulse response of the immobile zone is the same everywhere; (ii) there is no concentration gradient across the mobile zone except in the z direction; (iii) the flow rate is uniform in the mobile zone; (iv) mechanical dispersion is characterized by $D = \alpha V_m$.

Mass balance for advection, dispersion and an exchange between zones gives:

$$\theta_m \frac{\partial C_m^r}{\partial t} + \theta_{im} \frac{\partial \left\langle C_{im}^r \right\rangle}{\partial t} = \alpha V_m \theta_m \frac{\partial^2 C_m^r}{\partial z^2} - \theta_m V_m \frac{\partial C_m^r}{\partial z} \tag{5.1}$$

where C_m^r is the resident concentration of solute in the mobile zone, $\left\langle C_{im}^r \right\rangle$ is the average concentration in an immobile 'block', α is the dispersivity and V_m is the average advective fluid velocity in the mobile zone in the z direction. The r superscript indicates volume-

averaged or 'resident' concentration (see Appendix 2, which explains the relationship of resident and flux averaged concentrations).

Equation (5.1) needs to be accompanied by a second equation describing the exchange term, $\theta_{im}\frac{\partial\langle C^r_{im}\rangle}{\partial t}$. This exchange could be due to an advective flow at right-angles to the direction z, alternatively it might be due to a sorption kinetic. Very often it is assumed that the exchange is diffusively governed in which instance this second equation then becomes a statement of Fick's Second Law, with a particular geometry. In order to generalize, the nature of the exchange and the equation describing it are left unstated, for now.

First consider a general initial condition of $C^r_{m0}(z)$ and $\langle C^r_{im0}(z)\rangle$. The solution to Equation (5.1) can now be found by using the Laplace Transform (LT):

$$\bar{y}(s)=\int_0^\infty e^{-st}y(t)dt \tag{5.2}$$

where $y(t)$ is a function of t and s is the Laplace variable. Taking the LT of Equation (5.1) and rearranging, gives:

$$\frac{d^2\bar{C}^r_m}{dz^2}-\frac{1}{\alpha}\frac{d\bar{C}^r_m}{dz}-\frac{\bar{C}^r_m}{\alpha V_m}\underbrace{s\left(1+\sigma\frac{\langle\bar{C}^r_{im}\rangle}{\bar{C}^r_m}\right)}_{\mathrm{g}(s)}=\underbrace{-\frac{C_{m0}+\sigma\langle C^r_{im0}\rangle}{\alpha V_m}}_{\mathrm{I}(z)} \tag{5.3}$$

where $\sigma=\theta_{im}/\theta_m$. The solution to this second-order non-homogeneous ordinary differential equation is in general (e.g. Spiegel, 1968:105):

$$\bar{C}^r_m=Xe^{m_2z}+Ye^{m_1z}+\frac{1}{m_1-m_2}\left[e^{m_1z}\int_0^z \mathrm{I}(z)e^{-m_1u}du-e^{m_2z}\int_0^z \mathrm{I}(z)e^{-m_2u}du\right] \tag{5.4}$$

where X and Y are constants, $m_1=\frac{V_m}{2D}(1+\zeta)$, $m_2=\frac{V_m}{2D}(1-\zeta)$, $\zeta(s)=\sqrt{1+\frac{4\mathrm{g}(s)D}{V_m^2}}$,

$\mathrm{g}(s)=s\left(1+\sigma\,\mathrm{B}\left(\sqrt{st_{ch}}\right)\right)$ and $\mathrm{I}(z)=-\frac{\left(C_{m0}(z)+\sigma\langle C^r_{im0}(z)\rangle\right)}{\alpha V_m}$.

The Block Geometry Function[34] (BGF), B, can be defined as the response of the average concentration in an empty block to an impulsive change in concentration in the mobile zone,

$$\mathrm{B}=\frac{\langle\bar{C}^r_{im}\rangle}{\bar{C}^r_m} \tag{5.5}$$

[34] Barker (1985a&b) provided other definitions of the BGF, however this one is preferred here.

To establish X, Y, m_1, m_2 and I in Equation (5.4) requires the inlet and outlet boundary conditions and initial conditions to be specified.

5.1.2.1 Boundary and initial conditions

Simple standard boundary and initial conditions are established in Table 5.1 for the model comparison. A more detailed discussion of the effect of boundary conditions is contained in Appendix 2 (pages 443-459).

Table 5.1: Boundary and initial conditions (see Appendix 2 for definition of C^f and C^r, which are flux-averaged and volume-averaged concentrations respectively).

Condition	Options
Inlet boundary conditions (BCI) e.g. Impulse response	BCI(a) 'Type 1' : $C_0^r(0,t) = (m_0 / V_m)\delta(t)$ BCI(b) 'Type 3': $C_0^f(0,t) = \left(\bar{C}^r(0) - \alpha \left. \frac{\partial \bar{C}^r}{\partial z} \right\|_{z=0} \right) \delta(t) = (m_0 / V_m)\delta(t)$
Outlet boundary conditions (BCO)	BCO(a) Finite downstream boundary: $\left. \frac{dC^r}{dz} \right\|_L = 0$ BCO(b) Infinite downstream boundary: $\left. \frac{dC^r}{dz} \right\|_\infty = 0$
Initial condition (IC)	IC(0) Zero initial fracture and matrix resident concentration: $\langle \bar{C}_{im0}(z) \rangle = 0$ and $C_{m0}(z) = 0$ IC(1) Constant initial fracture and matrix resident concentration: $\langle \bar{C}_{im0}(z) \rangle = C_I$ and $C_{m0}(z) = C_I$

Note that M_0 is the total mass injected and $m_0 = M_0 / A$ is the mass injected per unit area. The input is parameterised as $(m_0 / V_m)\delta(t)$, in order to keep the solutions in terms of V_m. Impulse inputs are also commonly given in the literature as $(M_0 / Q)\delta(t)$. The latter is particularly useful for tracer experiments where a known mass, M_0, is suddenly added to a well which is injecting into an aquifer at Q.

If there is a constant (step) injection of $\dot{m}_0$ = (m_0 per unit time), then the input concentration will be $C_0 = \frac{\dot{m}}{V_m}$. Keeping this assumption ensures that conversion from the impulse to the step response simply requires an integral in time and conversion from the step to the impulse requires differentiation with respect to time.

Appropriate boundary conditions for 1D systems such as columns and soils are BCI(b) and BCO(b). These are now used as the boundary conditions for which different models are compared. The effect of changing the assumptions is discussed further in Appendix 2.

Hereafter the m sub-script is dropped and concentration is assumed to be mobile unless otherwise indicated.

Table 5.2 gives a summary of the lumped model framework.

5.1.3 Parameterisation

There is considerable choice over how to gather up the parameters in Equation (5.1) and its solutions for different conditions. A way that is used extensively here is assimilating in terms of characteristic times (e.g. $t_A = z/V$, $t_a = z/V_m$, $t_{cb} = b^2/D_a$, $t_D = \alpha/V$, $t_d = \alpha/V_m$, $t_{cf} = \frac{a^2 D_a}{4D_E^2}$). These could also be expressed as dimensionless times (e.g. $\gamma = t_a/t_{cb}$, or $Pe = t_A/t_D$). A difficulty with using normalized times or parameter groups is that large numbers can be generated when the terms in the denominator tend to zero. Which combination of parameters is optimum will depend on the task in hand. Parameterising the DP system with t_{cb} and σ (Fretwell, 1999; Ward et al., 1998; Wright, 2002) adds considerable clarity to the problem. This is because capturing key *timescales* is very useful. There is a difficulty getting meaningful data for fracture spacings, block size distributions or waste particle sizes. Additionally, it is possible that self-organisation of flow may create relatively immobile blocks which require more dynamic measurement. There is a further characteristic time, t_{cf}, which is of the order of time taken for solute to diffuse from a fracture into an empty block.[35] This is a useful way to parameterize systems, particularly in the presence of an effectively 'infinite' matrix when only t_{cf} (i.e. the ratio t_{cb}/σ^2) can be established. In the same way that using $t_{cb} = b^2/D_a$ can be used to relate block size to characteristic time, relating this time to dimensions is possible (using $t_{cf} = a^2 D_a/4D_E^2$). For systems approximated as having 1D vertical flow, such as in column experiments and landfills, there is a good chance of achieving well-constrained estimates of the total porosity (θ) and the total flux rate (q) through the system. One way of parameterising the DP system is using (t_a, t_{cb}, σ). t_a can be re-expressed as $z\theta/\left[q(1+\sigma)\right]$. If it can be assumed that θ, q and z are known, then there may be only two unknowns for a DP system, namely t_{cb} and σ.

[35] 'cb' stands for characteristic time for a block and 'cf' for characteristic time for a fracture.

Table 5.2: Summary of lumped models.

Model	Input function, $C^f(0,t)$	Laplace transform solution, $C^f(z,s)=(m_0/V_m)\bar{f}_\delta$ for impulse response and $C^f(z,s)=(\dot{m}_0/V_m)\bar{f}_H=C_0\bar{f}_H$ for step response	IC	BCI	BCO
AD- DP	$(m_0/V_m)\delta(t)$	$\bar{f}_\delta = Ae^{\frac{V_m z}{2D}(1-\zeta)} + Be^{\frac{V_m z}{2D}(1+\zeta)} + \frac{1}{m_1-m_2}\left[e^{m_1 z}\int_0^z I(z)e^{-m_1 u}du - e^{m_2 z}\int_0^z I(z)e^{-m_2 u}\right]$, where A and B are constants, $I=-\frac{C_{m0}+\sigma\langle C_{im0}\rangle}{\alpha V_m}$, $m_1=\frac{V_m}{2D}(1+\zeta)$, $m_2=\frac{V_m}{2D}(1-\zeta)$, $\zeta=\sqrt{1+\frac{4g(s)D}{V_m^2}}$, and $g(s)=s\left(1+\sigma B\left(\sqrt{st_{cb}}\right)\right)$	General	b	b
AD- DP	$(m_0/V_m)\delta(t)$	$\bar{f}_\delta = e^{\frac{V_m z}{2D}(1-\zeta)}$	IC(0)	b	b
DP	$(m_0/V_m)\delta(t)$	$\bar{f}_\delta = e^{-t_a g(s)}$	IC(0)	b	b
AD	$(m_0/V)\delta(t)$	$\bar{f}_\delta = e^{\frac{Vz}{2D}(1-\zeta)}$, where $\zeta=\sqrt{1+\frac{4sD}{V^2}}$	IC(0)	b	b
A (Piston)	$(m_0/V)\delta(t)$	$\bar{f}_\delta = e^{-\frac{z}{V}s}$, where D=0	IC(0)	b	b
AD & DP	$(m_0/V_m)\delta(t)$	$\bar{f}_\delta = \frac{4\zeta}{(1+\zeta)^2-(1-\zeta)^2\exp\left(-\frac{V_m L\zeta}{D}\right)}\exp\left[\frac{V_m L}{2D}(1-\zeta)\right]$	IC(0)	b	a
AD, DP, AD-DP or A	$C_0\mathrm{H}(t)$	$\bar{f}_H=\frac{C_0}{s}\bar{f}_\delta$, where $\bar{f}_\delta$ can be selected from above depending on BCs.	IC(0)	b	a or b
AD, DP, AD-DP or A	$0\mathrm{H}(t)$	$\bar{f}_H==\frac{C_I}{s}\left(1-\bar{f}_\delta\right)$, where $\bar{f}_\delta$ can be selected from above depending on BCs.	IC(1)	b	b

5.2 *Advection Dispersion model*

The behaviour, late-time gradients, moments and fitting of the AD model are now examined in turn. Unlike for DP and sorption processes, there are many and varied analytical solutions for the AD model in the time domain (as well as in Laplace space). Because of the existence of a dispersive flux across boundaries, the definition of concentration and boundary conditions are particularly important. In order to keep this chapter focussed on analysis of the 'standard' conditions by the three methods, the effects of boundary conditions are considered in Appendix 2 (pages 443-459).

5.2.1 Time-domain solutions

For the AD model the LT solution for Type 3 concentration input, BCI(b), and an infinite downstream boundary, BCO(b) (the standard boundary condition for this comparison) with an initially solute-free system, IC(0) is:

$$\bar{C}_\delta^f(s) = (m_0/V)\exp\left[\frac{Vz}{2D}\left(1-\sqrt{1+\frac{4sD}{V^2}}\right)\right] \tag{5.6}$$

This inverts to:

$$C_\delta^f = (m_0/V)\frac{z}{\sqrt{4\pi Dt^3}}\exp\left[-\frac{(z-Vt)^2}{4Dt}\right] \tag{5.7}$$

The solution for a Type 1 flux input and infinite downstream boundary with a step input gives:

$$\bar{C}_\delta^f(s) = \frac{C_0}{s}\exp\left[\frac{Vz}{2D}\left(1-\sqrt{1+\frac{4sD}{V^2}}\right)\right] \tag{5.8}$$

This inverts to give the 'Ogata Banks' equation (Equation 3.10, with $R = 1$).

A 'top hat' response in the time domain can be readily obtained by linear superposition of two step responses.

5.2.2 Late-time gradients

5.2.2.1 Impulse response

The late-time gradient is calculated for Equation (5.7). Taking logs of both sides gives

$$\ln(C) = \ln\left(\frac{m_0 z}{\sqrt{4\pi\alpha V^3}}\right) - \frac{3}{2}\ln(t) - \frac{(z-Vt)^2}{4\alpha Vt} \tag{5.9}$$

Differentiating with respect to t gives,

$$\frac{\partial}{\partial t}\ln(C) = \left(\frac{z^2}{4\alpha V}\right)\frac{1}{t^2} - \left(\frac{3}{2}\right)\frac{1}{t} - \frac{V}{4\alpha} \tag{5.10}$$

The first term for $\frac{\partial}{\partial t}\ln(C)$ dominates when t is small (i.e. $t \ll \frac{z^2}{6\alpha V}$), but becomes insignificant for large t. The second term is an intermediate condition of a log-log gradient of -3/2 in the instance that the first and third terms are small (i.e. $\frac{z^2}{6\alpha V} \ll t \ll \frac{6\alpha}{V}$). See Figure 5.3. For long enough time the first and second terms become negligible and the third term gives a log-lin linear gradient of $-V/4\alpha$.

The log-log linear gradient of -3/2 is not a characteristic of the AD system. One reason for this mistake may be that the scale effect (Section 3.2.1.1.2) gives an approximate correlation of $\alpha = 0.1z$ or $Pe = 10$. Applying this to the criterion for the -3/2 gradient section requires $\frac{\alpha}{z} \gg \frac{1}{6}$ or $Pe \ll 6$, which is a rather small Pe compared to what is normally measured. So an AD log-log gradient of -3/2 is mathematically possible, but physically unlikely.

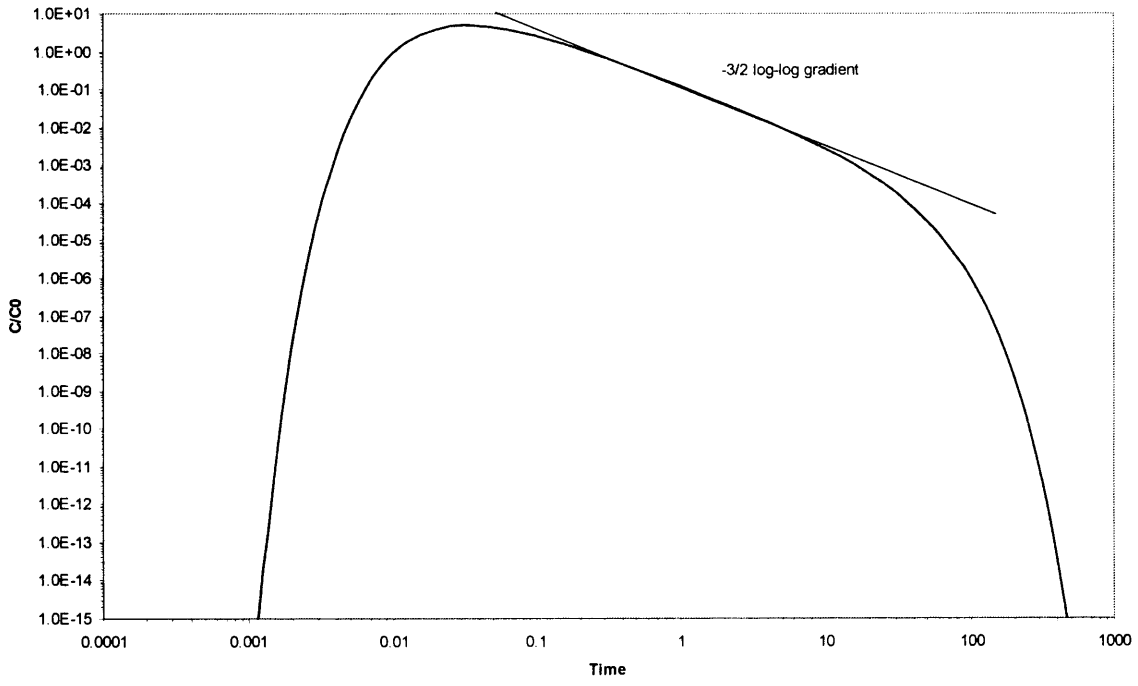


Figure 5.3: Period of log-log linear -3/2 gradient for the AD model with $z = 1$, $\alpha = 5$ and $V = 1$.

5.2.2.2 Flushing response

The flushing response is:

$$C_F = C_I - C_I \int_0^t \frac{z}{\sqrt{4\pi\alpha V t^3}} e^{-\frac{(z-Vt)^2}{4\alpha V t}} dt \qquad (5.11)$$

The gradient is given by:

$$\frac{\partial C_F}{\partial t} = -G_\delta = -C_I \frac{z}{\sqrt{4\pi\alpha V t^3}} \exp\left[-\frac{(z-Vt)^2}{4\alpha V t}\right] \tag{5.12}$$

This is compared to the gradient of the impulse response

$$\frac{\partial C_\delta}{\partial t} = \frac{z e^{-\frac{(z-Vt)^2}{4\alpha V t}}}{8\sqrt{\pi}\alpha^2 t^5 V^2}\left[z^2 - Vt(6\alpha + Vt)\right]\sqrt{aVt^3} \tag{5.13}$$

which for $t \gg \frac{\alpha}{V}$ simplifies to

$$\frac{\partial C_\delta}{\partial t} = -\left(\frac{z}{\sqrt{4\pi\alpha V t^3}} \exp\left[-\frac{(z-Vt)^2}{4\alpha V t}\right]\right)\left(\frac{6\alpha + Vt}{4\alpha t}\right) \tag{5.14}$$

For $t \gg \frac{z}{V}$ the gradients of the impulse response and flushing system therefore differ by a factor of $\frac{V}{4\alpha}$.

But, as these gradients become very small (tends to zero as t tends to infinity) the absolute difference between them becomes very small too. So, in practise the gradients of $\ln(C)$ against t converge. So the very late time gradient of the flushing response is also approximately exponential.

5.2.3 Moments

The first few moments are compared, where $t_A = \frac{z}{V}$ and $t_D = \frac{\alpha}{V}$:

$$\mu'_1 = t_A \tag{5.15}$$

$$\mu_2 = 2t_A t_D \tag{5.16}$$

$$\mu_3 = 12 t_A t_D^2 \tag{5.17}$$

$$\mu_4 = 12 t_A t_D^2 (t_A + 10 t_D) \tag{5.18}$$

5.2.4 Fitting of the AD model

5.2.4.1 General observations of the AD model BTC

The behaviour of the impulse response BTC is a little surprising with respect to the change that occurs when α is varied (Figure 5.4). There is a point of a minimum peak. This means that the peak concentration can increase as α is increased, which is not intuitive. Because the effect does not occur for the response to an initial $\delta(z)$, this is probably an effect due to the inlet boundary condition.

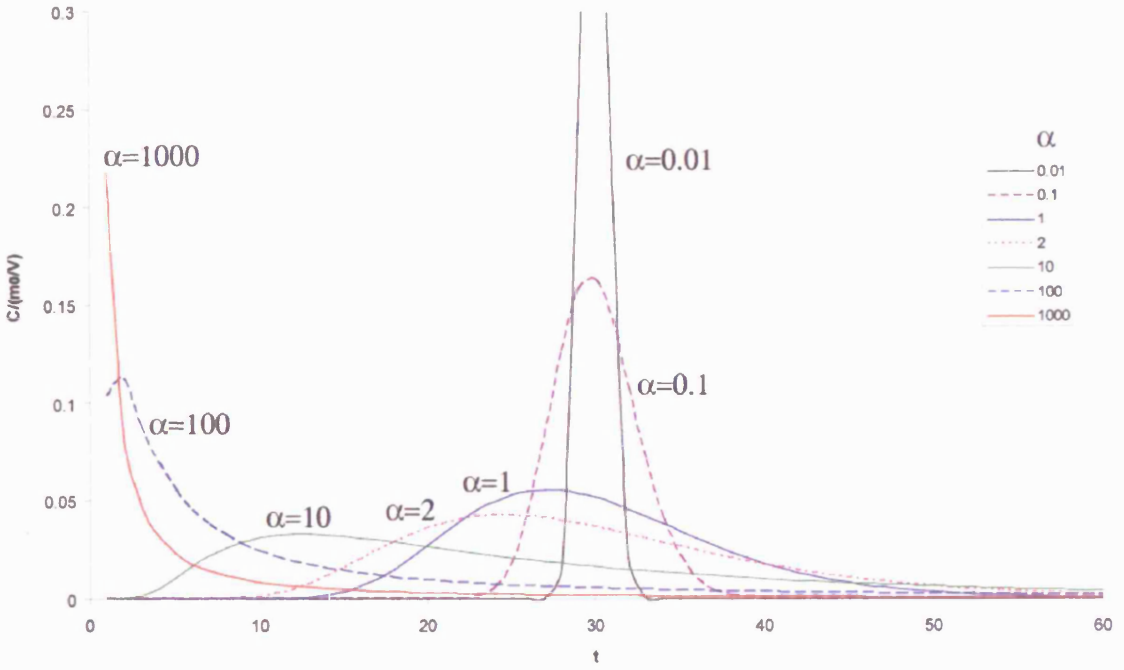


Figure 5.4: AD with $V = 1$, $z = 30$.

5.2.4.2 Monte Carlo method applied to the Ad model with 'Approach 1'

A Monte Carlo (MC) analysis was applied to the AD impulse-response (Equation 5.7) fit to synthetic data from the same model. The aim was to explore the shape of the confidence interval and assess whether there is much effect from choice of parameter value, parameterisation and noise.

An example is given here in Figure 5.5, where 100 points on the BTC are compared (equally spaced by times step of 1). 9529 out of 10,000 points lay on or inside the ellipses. 95% confidence intervals were found to be $V_{\text{upper}} = 2.05$, $V_{\text{lower}} = 1.95$, $\alpha_{\text{upper}} = 2.11$, $\alpha_{\text{lower}} = 1.89$.

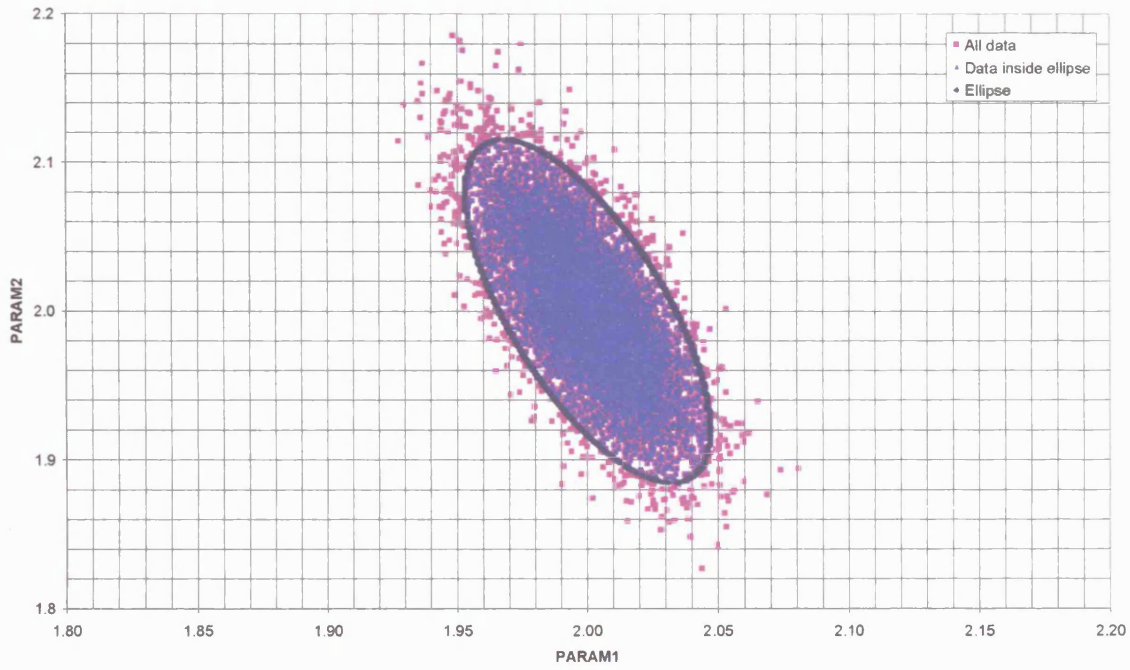


Figure 5.5: MC of AD model against AD-synthetic data (PARAM1=V , PARAM2=α). The base BTC is generated with $z = 10$, $V = 2$, $\alpha = 2$ and Gaussian noise with $\sigma_N = 0.01(m_0 / V)$.

The data appeared to form a reasonably fat and smooth ellipse and the 95% confidence interval was accurately approximated.

This process was repeated for the complete range of parameterisation possibilities as given in Section 5.1.4. It was found that parameterisation makes no difference to the shape of the ellipse, and scales the axes accordingly.

There was a very good estimation of the ellipse as obtained by the Hessian method (Press et al., 1992) for a wide range of base parameters (there was always a very small error in the number counted in or on the ellipse). The only instance of divergence of the estimate to what was counted numerically was in the instance of a very short t_a and when $\sigma_N > 0.1C_0$ (which is an unrealistically noisy signal).

5.3 *Dual-Porosity (in absence of mechanical dispersion)*

Having considered the AD process in isolation from DP or sorption effects, this section now examines the DP process in isolation. Recall, the rationale here is to better diagnose processes, in pursuit of objective 2: *to develop a comparative framework of a sub-set of appropriate models for diagnosis of simple processes.*

The DP model is now examined using late-time gradients, method of moments and by fitting.

The effect of changing t_{cb}, whilst holding the other parameters constant, is shown in Figure 5.6 and on log-lin axes in Figure 5.7. For $t \ll t_{cb}$ there is effectively an infinite matrix. How an infinite matrix affects the results depends on the time taken to diffuse from the mobile zone, t_{cf}, relative to the advection time, t_a. If $t_{cf} \gg t_a$, then very little of the solute in the fracture will have had chance to leave the fracture by the time the mobile pulse reaches the outlet. For short time on linear axes the BTC looks like an undispersed pulse that only passes through the mobile porosity (Figure 5.6). However, plotting on log axes reveals a very long tail that is not visible on linear axes (Figure 5.7). For increasing values of t, for $t \geq t_{cb}$, the immobile phase is closer to being in equilibrium with the mobile phase and the system tends to another undispersed pulse, but one which characterises the total porosity (therefore arriving later). Viewed on normal axes it therefore appears that there is an intermediate value of t_{cb} (which is of the order of t_a), which gives the most spread out looking BTC. Increasing or decreasing t_{cb} apparently causes a tighter arrival. This will be shown to be not borne out by moment analysis in Section 5.3.3, and the reason for this is clear from the log-lin axes of Figure 5.7. Long t_{cb} will have a large second moment because of the contribution of the tail to the measurement of spread, rather than the contribution of the BTC that is visible on linear axes.

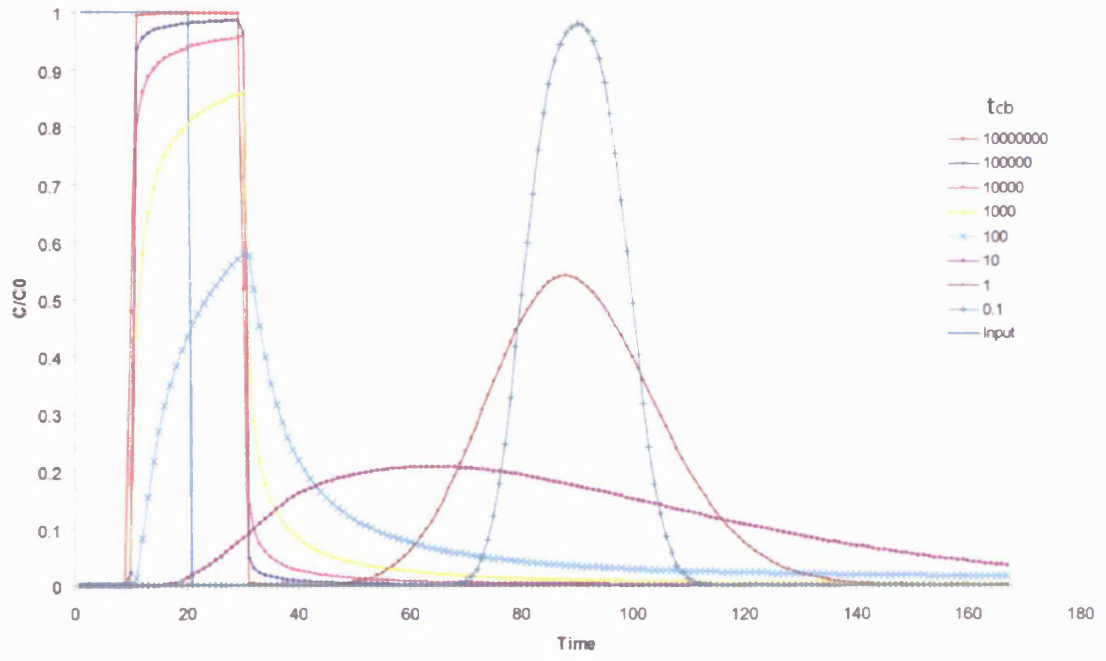


Figure 5.6: BTCs for DP model - effect of changing t_{cb} for $\sigma = 7$, t_a=10 for 'top-hat' input (input time, $T_p = 20$)

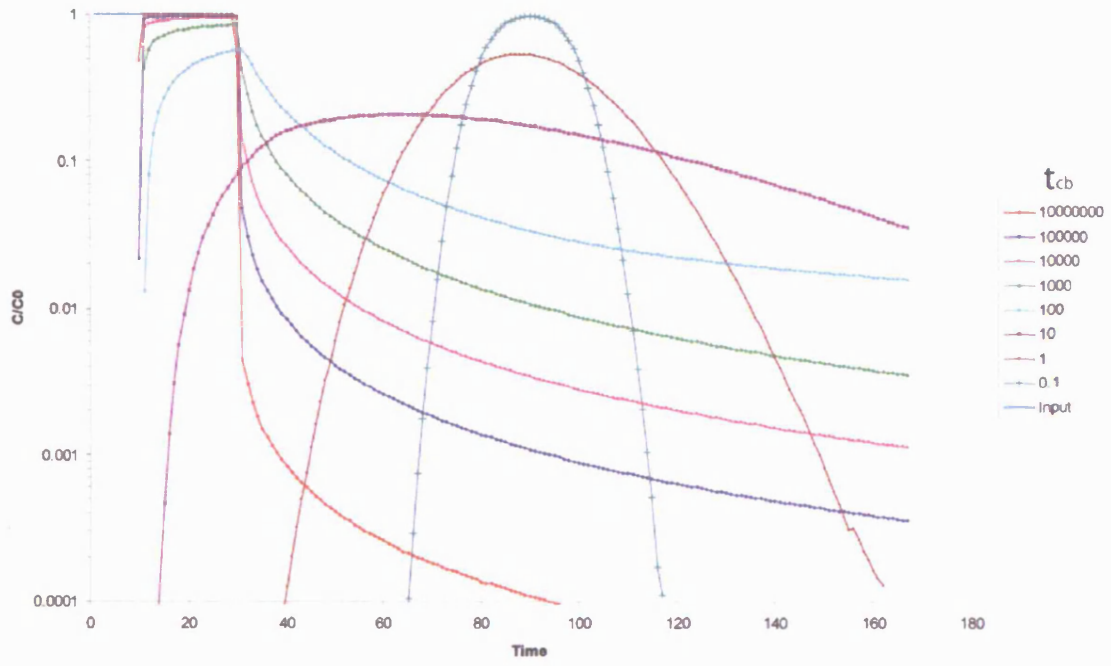


Figure 5.7: BTCs for DP model - effect of changing t_{cb} for $\sigma = 7$, t_a =10 for 'top-hat' input (input time, $T_p = 20$). Log-lin scale.

5.3.1 Block Geometry Functions

There are several equivalent definitions for the BGF, B, a number of which are detailed in Barker (1985a&b). Here the dimensionless BGF is defined in Laplace space as the impulse response of the average matrix concentration $\mathrm{B}(\sqrt{S}) = \dfrac{\langle \overline{C}_{im}(S) \rangle}{\overline{C}_m(S)}$, where the dimensionless Laplace variable is $S = st_{cb}$. Inverting gives $\mathrm{B}_t(T) = \dfrac{C_{im}(T)}{C_m(T)}$, where $T = t/t_{cb}$. Note that this is precisely the same as defining the BGF to be the flux into unit volume, Q, due to a step in solute concentration outside an initially empty block:

$$\overline{Q}_{inH} = S\frac{\mathrm{B}(\sqrt{S})}{S} = \mathrm{B}(\sqrt{S}) = \langle \overline{C}_{im\delta} \rangle \tag{5.19}$$

The impulse response of a block (i.e. to $C_0\delta(t)$ at the edge of the block) depends on the geometry of the block. Since there is a large body of work on solving the diffusion (heat) equation for different boundary conditions, there are a number of analytical solutions available for the BGF. Some of these are summarized in Barker (1985a) and Carneiro (2000). The most applicable and widely cited of these that are of relevance here are the many solutions contained within Carslaw and Jaeger (1959) and Crank (1975).

A typical step solution is given for the slab Carslaw and Jaeger (1959:83) and is shown here reparameterised:

$$\left.\frac{C_{im}(Z,T)}{C_0}\right|_H = 1-2\sum_{n=0}^{\infty}\frac{(-1)^n}{A_n^{1/2}}e^{-A_nT}\cos(A_n^{1/2}Z) \tag{5.20}$$

where $Z=\frac{z}{b}$ and $A_n=(2n+1)^2\pi^2/4$

Differentiate w.r.t. T to give the impulse response:

$$\left.\frac{C_{im}(Z,T)}{C_0}\right|_\delta = 2\sum_{n=0}^{\infty}(-1)^n A_n^{1/2}e^{-A_nT}\cos(A_n^{1/2}Z) \tag{5.21}$$

Integrate over Z to give the average concentration:

$$\left.\frac{\langle C_{im}(T)\rangle}{C_0}\right|_\delta = 2\sum_{n=0}^{\infty}(-1)^n e^{-A_nT}\left[\sin(A_n^{1/2}Z)\right]_0^1 = 2\sum_{n=0}^{\infty}e^{-A_nT} \tag{5.22}$$

Compare this to the LT solution for the slab:

$$\left.\frac{\langle \overline{C}_{im}(S)\rangle}{C_0}\right|_\delta = \sum_{n=0}^{\infty}\frac{2}{S+A_n} = \frac{\tanh\left(\sqrt{S}\right)}{\sqrt{S}} = \mathrm{B}\left(\sqrt{S}\right) \tag{5.23}$$

It is useful to examine the BGF in both the time-domain and in Laplace space. In the time domain the later terms in the infinite series of exponentials are of diminishing importance at long time. For very long time the first component (the smallest eigenvalue) dominates.

The BGFs for common simple geometries (such as for the slab, sphere and cylinder) appear under many different names in the literature. Here just a few publications in which the BGF appears are considered in Table 5.3.

Table 5.3: Different approaches to the BGF

Author	**Representation for BGF,** $\mathrm{B}\left(\sqrt{s}\right)$
Cunningham and Roberts, 1998	$\mathrm{W}(s)$
Carrera et al., 1998	$\mathrm{g}(s)/\sigma$
Haggerty et al., 2000	$\mathrm{g}(s)/\sigma$
Herrera, 1976	$\mathrm{g}(s)/\sigma$
Sardin et al., 1991	$\mathrm{M}(s)$, $\mathrm{H}^*(s)$
Haggerty and Gorelick, 1995	Infinite sum of MIM exchanges
Kazemi, 1969; Aris, 1975; van Genuchten and Dalton, 1986	Shape factors
Carslaw and Jaeger, 1959	Impulse responses for different geometries

The BGF appears even more often, but in a less general way and buried in solutions that are specific to particular geometries. Note that of course it also appears in literature on pressure

transient behaviour as well as that specifically for solutes. Figure 5.8 and Figure 5.9 show typical BGFs.

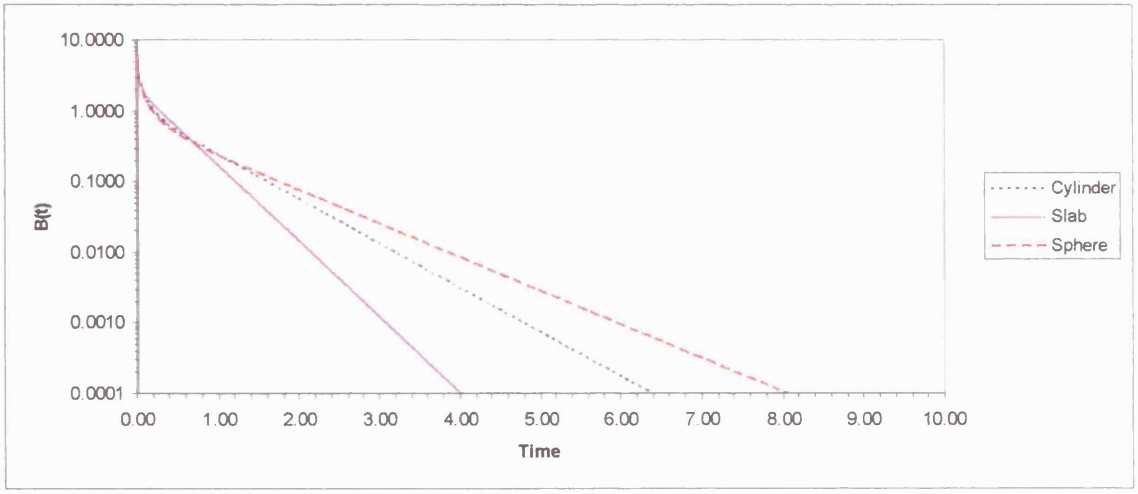


Figure 5.8: Log-lin BGFs, $B_t(t)$ for cylinder, slab and sphere – gradients are 1.45. $\pi^2/4$ and $\pi^2/9$, respectively.

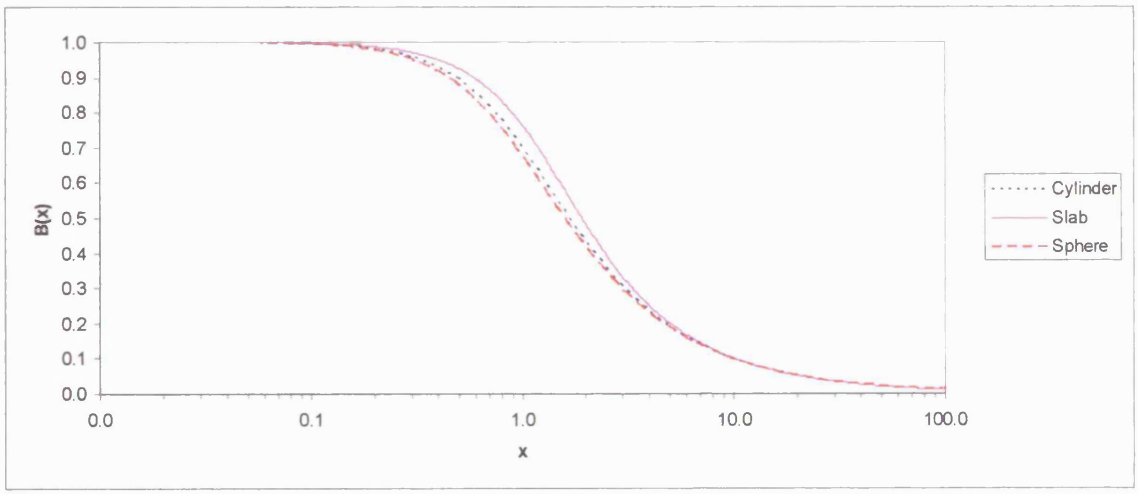


Figure 5.9: Log-lin BGF, $B(x)$ for cylinder, slab and sphere.

The BGF for the slab, cylinder and sphere can all be expressed as a single function (e.g Sanchez-Vila and Carrera (2004)) each distinguished by parameter, μ, where μ is 1,2,3 for slabs, cylinders and spheres respectively. The expression is given as $B(x) = \frac{1}{x}\frac{I_{\mu/2}(\mu x)}{I_{\mu/2-1}(\mu x)}$ (Barker, 1985b), where I_z is a modified Bessel function of the first kind of order, z. μ can hold non-integer values, which when between 1 and 3 representing intermediate shapes between the three simple geometries. This representation gives a fairly tight family of curves which are

only a subset of the possible BGFs that might exist in nature. As the sphere dimensions tends to infinity the BGF tends to $\mathrm{B}(x)=\dfrac{2}{1+\sqrt{1+4x^2}}$ (Barker, personal communication).

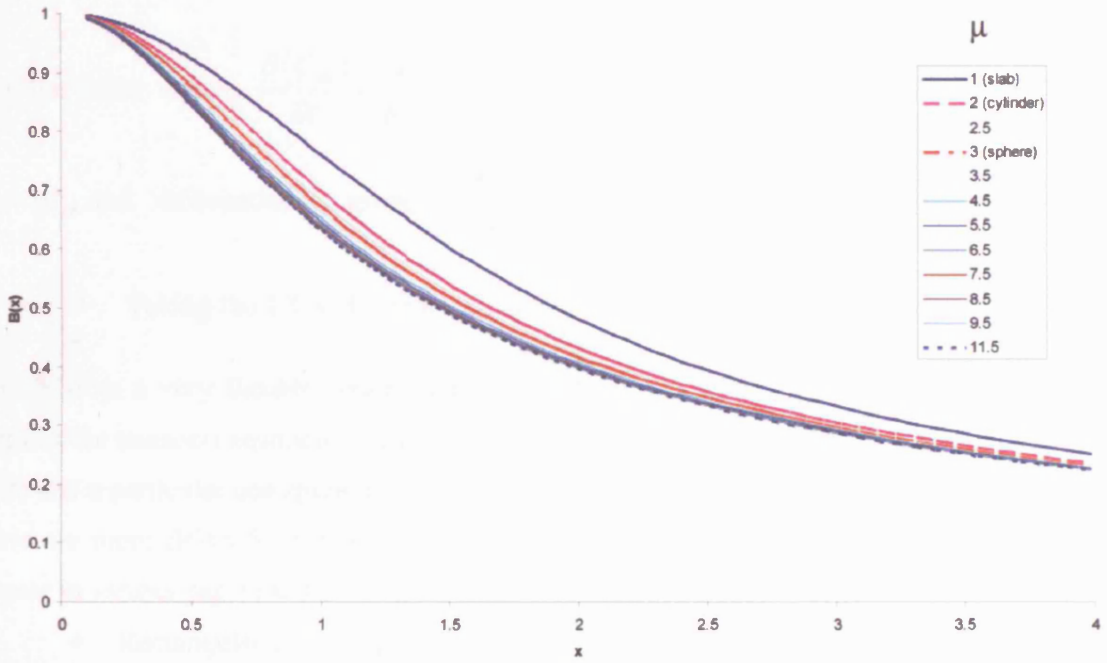


Figure 5.10: Family of BGFs, $\mathrm{B}(x)$ parameterized only by μ (x-axis compressed to enhance the difference between the curves).

Note that these common BGFs share asymptotic limits[36]: $\mathrm{B}(\sqrt{S})\big|_{S\to 0}=1$, $\mathrm{B}(\sqrt{S})\big|_{S\to\infty}=0$, and the condition that $\dfrac{d}{dS}\mathrm{B}(\sqrt{S})\le 0$. Since $\tanh\left(\sqrt{S}\right)$ tends to 1 as $\sqrt{S}$ tends to infinity, $\tanh\left(\sqrt{S}\right)/\sqrt{S}\sim 1/\sqrt{S}$. This also happens for other geometries, for example for a sphere $\mathrm{B}\left(\sqrt{S}\right)=\coth\left(3\sqrt{S}\right)/\sqrt{S}-1/(3S)\sim 1/\sqrt{S}$.

Inverting the LT to return to the time domain, $L^{-1}\dfrac{1}{\sqrt{S}}=\dfrac{1}{\sqrt{\pi T}}$ (Abramowitz and Stegun, 1968) and therefore:

$$\lim_{T\to 0}\left[\mathrm{B_t}(T)\right]=\frac{1}{\sqrt{\pi T}} \tag{5.24}$$

[36] To what extent this is completely generic is an interesting question. It would be useful to formally define what conditions are required of a block in order that the asymptotes of the BGF are always 0 and 1.

Note that step response for infinite block: $C_{im}(z,t)=C_0\text{erfc}\left(\frac{z}{2\sqrt{t}}\right)$. The step response for diffusion into a block is $\frac{\partial C_{im}}{\partial t}=\frac{\partial}{\partial t}C_0\text{erfc}\left(\frac{z}{2\sqrt{t}}\right)=-\frac{C_0}{\sqrt{\pi t}}$.

For the MIM model, $\frac{\partial\langle C_{im}\rangle}{\partial t}=\frac{\kappa}{\theta_{im}}(C_m-\langle C_{im}\rangle)$. Expressing this in dimensionless time, $T=t/t_{cb}$ and differentiating gives, $\frac{\partial\langle C_{im}\rangle}{\partial T}=\frac{\alpha t_{cb}}{\theta_{im}}\left(C_m-\langle C_{im}\rangle\right)=\kappa'\left(C_m-\langle C_{im}\rangle\right)$, where $\kappa'=\frac{\kappa t_{cb}}{\theta_{im}}$. Taking the LT and rearranging gives the BGF, $B(\sqrt{S})$ for the MIM as $\frac{\kappa'}{S+\kappa'}$.

The BGF is a very flexible concept allowing different geometrical assumptions to be tested without the transport equations having to be completely recast. Several BGFs can be stored in a code and a particular one specified in an input file. Beyond simple slabs, cylinders and spheres there are more BGFs for which analytical solutions have been derived by Barker and which appear in various papers and PhD theses. Some examples follow:

- Rectangular parellepiped (Barker, 1985a), with the cube as a special case (Barker, 1985b);
- finite cylinder (Carneiro, 2000);
- hollow cylinder (Carneiro, 2000);
- hierarchical BGF (Carneiro, 2000);
- n-dimensional spheres (Barker, 1988 & 1985b; Carneiro, 2000; Wright, 2002);
- mixture of different block shapes Barker (1985a&b) (these have also been addressed by Neretnieks and Ramuson, 1984; Pruess and Narasimhan, 1985; Zimmerman and Bodvarsson, 1990).

The parellepiped, finite cylinder and hollow cylinder require infinite sums which in practice would need to be truncated. The precise geometrical interpretation of an N-dimensional sphere is difficult to conceptualize or relate to reality, however, the generalization of dimensionality turns out to be a very powerful tool (Barker, 1988). Part III considers the implication of interpretting different flow systems in a channel network model with different assumptions of flow dimension. A hybrid BGF comprised of mixed block shapes is formed by a simple summation (Barker, 1985b). Similarly mixtures of different block sizes can be obtained by summation or integration (see Chapter 7).

The justification for use of a particular BGF is ultimately limited by the knowability of the true geometry. For anything other than simple shapes the BGF may become analytically intractable. Tables 5.4 and 5.5 show some simple BGFs.

Table 5.4: Some simple BGFs, where $t_{cb} = b^2/D_a$ and b is block volume/area, u_j is the jth solution of $J_0(u_j) = 0$, where J_0 is a Bessel function of the first kind, $A_n = (2n+1)^2\pi^2/4$ and $\kappa' = \kappa t_{cb}/\theta_{im}$

Geom	$B_t(t)$	$B(\sqrt{st_{cb}})$ As functions	$B(\sqrt{st_{cb}})$ As series	Impulse response in concentration	Impulse response in flow $\Gamma(t) = \frac{\partial \langle C_{im} \rangle}{\partial t}$	Step response in concentration	Step response in flow $\Gamma(t) = \frac{\partial \langle C_{im} \rangle}{\partial t}$
Slab $b = h$	$2\sum_{n=0}^{\infty}\frac{1}{t_{cb}}e^{-A_n t/t_{cb}}$	$\tanh(\sqrt{st_{cb}})/\sqrt{st_{cb}}$	$\sum_{n=0}^{\infty}\frac{2}{st_{cb}+A_n}$	$2\sum_{n=0}^{\infty}\frac{1}{t_{cb}}e^{-A_n t/t_{cb}}$	$-2\sum_{n=0}^{\infty}\frac{A_n}{t_{cb}}e^{-A_n t/t_{cb}}$	$1-2\sum_{n=0}^{\infty}\frac{1}{A_n}e^{-A_n t/t_{cb}}$	$2\sum_{n=0}^{\infty}\frac{1}{t_{cb}}e^{-A_n t/t_{cb}}$
Cylinder $b = r/2$	$2\sum_{n=1}^{\infty}\frac{1}{t_{cb}}e^{-u_j^2 t/t_{cb}}$	$\frac{I_1(2\sqrt{st_{cb}})}{\sqrt{st_{cb}}I_o(2\sqrt{st_{cb}})}$	$\sum_{n=1}^{\infty}\frac{4}{st_{cb}+u_j^2}$	$\sum_{n=1}^{\infty}\frac{4}{t_{cb}}e^{-u_j^2 t/t_{cb}}$	$-4\sum_{n=1}^{\infty}u_j^2 e^{-u_j^2 t/t_{cb}}$	$1-4\sum_{n=1}^{\infty}u_j^2 e^{-u_j^2 t/t_{cb}}$	$4\sum_{n=1}^{\infty}\frac{1}{t_{cb}}e^{-u_j^2 t/t_{cb}}$
Sphere $b = r/3$	$2\sum_{n=1}^{\infty}\frac{1}{t_{cb}}e^{-n^2\pi^2 t/t_{cb}}$	$\coth\left(3\sqrt{st_{cb}}\right)/\sqrt{st_{cb}} - 1/(3st_{cb})$	$\sum_{n=1}^{\infty}\frac{6}{st_{cb}+n^2\pi^2}$	$6\sum_{n=1}^{\infty}\frac{1}{t_{cb}}e^{-n^2\pi^2 t/t_{cb}}$	$-6\sum_{n=1}^{\infty}\frac{n^2\pi^2}{t_{cb}}e^{-n^2\pi^2 t/t_{cb}}$	$1-6\sum_{n=1}^{\infty}\frac{1}{n^2\pi^2}e^{-n^2\pi^2 t/t_{cb}}$	$6\sum_{n=1}^{\infty}\frac{1}{t_{cb}}e^{-n^2\pi^2 t/t_{cb}}$
MIM -	$\frac{\kappa}{\theta_{im}}e^{-\kappa t/\theta_{im}}$	$\frac{\kappa}{s\theta_{im}+\kappa}$	$\frac{\kappa}{s\theta_{im}+\kappa}$	$\frac{\kappa}{\theta_{im}}e^{-\kappa t/\theta_{im}}$	$-\frac{\kappa^2}{\theta_{im}^2}e^{-\kappa t/t_{cb}}$	$1-e^{-\kappa t/\theta_{im}}$	$\frac{\kappa}{\theta_{im}}e^{-\kappa t/\theta_{im}}$

Table 5.5: Some simple BGFs (dimensionless time), where $t_{cb} = b^2/D_a$ and b is block volume/area, u_j is the jth solution of $J_0(u_j) = 0$, where J_0 is a Bessel function of the first kind, $A_n = (2n+1)^2\pi^2/4$ and $\kappa' = \kappa t_{cb}/\theta_{im}$.

Geom	$B_t(T)$	$B(\sqrt{S})$ **As functions**	$B(\sqrt{S})$ **As series**	**Impulse response in concentration**	**Impulse response in flow** $\Gamma(T) = \frac{\partial \langle C_{im} \rangle}{\partial T}$	**Step response in concentration**	**Step response in flow** $\Gamma(T) = \frac{\partial \langle C_{im} \rangle}{\partial T}$
Slab $b = h$	$2\sum_{n=0}^{\infty} e^{-A_n T}$	$\tanh(\sqrt{S})/\sqrt{S}$	$\sum_{n=0}^{\infty} \frac{2}{S + A_n}$	$2\sum_{n=0}^{\infty} e^{-A_n T}$	$-2\sum_{n=0}^{\infty} A_n e^{-A_n T}$	$1 - 2\sum_{n=0}^{\infty} \frac{1}{A_n} e^{-A_n T}$	$2\sum_{n=0}^{\infty} e^{-A_n T}$
Cylinder $b = r/2$	$2\sum_{n=1}^{\infty} e^{-u_j^2 T}$	$\frac{I_1(2\sqrt{S})}{\sqrt{s} I_o\left(2\sqrt{S}\right)}$	$\sum_{n=1}^{\infty} \frac{4}{S + u_j^2}$	$4\sum_{n=1}^{\infty} e^{-u_j^2 T}$	$-4\sum_{n=1}^{\infty} u_j^2 e^{-u_j^2 T}$	$1 - 4\sum_{n=1}^{\infty} u_j^2 e^{-u_j^2 T}$	$4\sum_{n=1}^{\infty} e^{-u_j^2 T}$
Sphere $b = r/3$	$2\sum_{n=1}^{\infty} e^{-n^2\pi^2 T}$	$\coth(3\sqrt{S})/\sqrt{S} - 1/(3S)$	$6\sum_{n=1}^{\infty} \frac{1}{S + n^2\pi^2}$	$6\sum_{n=1}^{\infty} e^{-n^2\pi^2 T}$	$-6\sum_{n=1}^{\infty} n^2\pi^2 e^{-n^2\pi^2 T}$	$1 - 6\sum_{n=1}^{\infty} \frac{1}{n^2\pi^2} e^{-n^2\pi^2 T}$	$6\sum_{n=1}^{\infty} e^{-n^2\pi^2 T}$
MIM -	$\kappa' e^{-\kappa' T}$	$\frac{\kappa'}{S + \kappa'}$	$\frac{\kappa'}{S + \kappa'}$	$\kappa' e^{-\kappa' T}$	$-\kappa'^2 e^{-\kappa' T}$	$1 - e^{-\kappa' T}$	$\kappa' e^{-\kappa' T}$

5.3.2 Late-time gradients of the DP model

Haggerty et al. (2000) approached examination of the late-time gradient of a BTC in an AD-DP system by expressing the exchange term as a memory function in the time-domain (which is reparameterised here in terms of the BGF). Rearranging Equation (5.1) for the exchange flux in the AD-DP model gives:

$$\sigma\frac{\partial\langle C_{im}\rangle}{\partial t}=\alpha V_m\frac{\partial^2 C_m}{\partial z^2}-V_m\frac{\partial C_m}{\partial z}-\frac{\partial C_m}{\partial t}$$

Haggerty et al. (2000) considered the situation of late-time (i.e. $t\gg t_a$) and made two key approximations:

- The advective flux dominates over the dispersive flux, i.e. $\alpha V_m\frac{\partial C_m}{\partial z}\ll C_m V_m$.
- The immobile zone has a long residence time compared to advection, i.e. $t_{cb}\gg t_a$, so that the exchange term dominates the concentration in the mobile zone, i.e. $\sigma\frac{\partial\langle C_{im}\rangle}{\partial t}\gg\frac{\partial C_m}{\partial t}$.

After some algebra, the key result (in BGF nomenclature) is,

$$C_m(t_a,t)\approx t_a\left(C_I\sigma \mathrm{B}_t(t)-(m_0/V_m)\sigma\frac{d\mathrm{B}_t(t)}{dt}\right) \tag{5.25}$$

This is readily simplified to give the impulse response (when $C_I=0$)

$$C_m(t_a,t)\approx -t_a(m_0/V_m)\sigma\frac{d\mathrm{B}_t(t)}{dt} \tag{5.26}$$

Taking LTs gives the impulse response in Laplace space:

$$\bar{C}_m(t_a,s)\approx -(m_0/V_m)t_a\sigma s\mathrm{B}\left(\sqrt{st_{cb}}\right) \tag{5.27}$$

The flushing response (i.e. $C_0(0,t)=0$ and $C(z,0)=C_I$) is given by,

$$C_m(t_a,t)\approx t_a C_I\sigma \mathrm{B}_t(t) \tag{5.28}$$

Taking LTs gives the flushing response in Laplace space:

$$\bar{C}_m(t_a,s)\approx C_I t_a\sigma\mathrm{B}\left(\sqrt{st_{cb}}\right) \tag{5.29}$$

This approximation is extremely valuable. It allows the late time mobile and immobile concentrations or gradients to be estimated without requiring a laborious numerical solution. Haggerty et al. (2000) focussed on expressing the memory function in terms of infinite sums of MIM exchanges, but there is no motivation for doing that here.

The time-domain BGFs can be expressed as infinite sums of exponentials. A log-lin linear response at very long time is therefore predicted. The log-log gradient however, is unlikely to have any constant period, except for when $t \ll t_{cb}$ and the matrix is effectively infinite.

The infinite matrix response is given by,

$$C_\delta = (m_0 / V_m)\frac{t_a\sigma}{\sqrt{4\pi t_{cb}(t-t_a)^3}}\exp\left[-\frac{t_a^2\sigma^2}{4t_{cb}(t-t_a)}\right] \tag{5.30}$$

The gradient is achieved by taking logs of both sides and differentiating, to give:

$$\frac{\partial}{\partial t}\ln(C) = \frac{t_a^2\sigma^2 + 6t_a t_{cb} - 6t_{cb}t}{4t_{cb}(t_a - t)^2} \tag{5.31}$$

For $t \gg t_a$,

$$\frac{\partial}{\partial t}\ln(C) = \frac{t_a^2\sigma^2}{4t_{cb}t^2} + \frac{6t_a}{4t^2} - \frac{3}{2t} = \frac{t_a^2\sigma^2}{4t_{cb}t^2} + \frac{3t_a}{2t^2} - \frac{3}{2t} \tag{5.32}$$

Therefore the last term gives a log-log gradient of -3/2 when $t \gg \frac{t_a^2\sigma^2 + 6t_a t_{cb}}{6t_{cb}}$. This expression, incidentally, gives the peak time.

For long-time, $t \gg t_a$, so Equation (5.30) simplifies to:

$$C = \frac{t_a\sigma}{\sqrt{4\pi t_{cb}t^3}}\exp\left[-\frac{t_a^2\sigma^2}{4t_{cb}t}\right] \tag{5.33}$$

Taking logs gives:

$$\ln(C) = \ln(t_a) + \ln\left(\frac{\sigma}{\sqrt{4\pi t_{cb}t^3}}\right) - \frac{t_a^2\sigma^2}{4t_{cb}t} \tag{5.34}$$

where the last term is vanishingly small, so

$$\ln(C) \approx -\frac{3}{2}\ln(t) + \underbrace{\ln(\sigma) + \ln(t_a) - \ln\left(\sqrt{4\pi t_{cb}}\right)}_{\text{Offset}} \tag{5.35}$$

So $\ln(C)$ against $\ln(t)$ will have a gradient of -3/2 and a vertical offset that is determined by $\ln(\sigma)$, $\ln(t_a)$ and $\ln\left(\sqrt{4\pi t_{cb}}\right)$. If all parameters are constant and σ is changed, the offset will vary by $\ln(\sigma)$. This might arise, for example, when different sized solutes that can access different porosities are used in a multi-tracer experiment. Table 5.6 summaries the late time gradients for the DP model.

Table 5.6: Periods of constant log-log or log-lin gradient.

Process	log-log	log-lin
DP	-3/2 when t<<t_{cb}	First eigenvalue, e.g. : $-\pi^2/4$ slab $-\pi^2/9$ sphere -1.45 cylinder
MIM	Never	$-\kappa/\theta_{im}$

The MIM log-lin gradient can be matched to the DP for very long time ($t \gg t_{cb}$). For example, for a slab, $\kappa = \theta_{im}\pi^2/4$.

Haggerty et al. (2000) extended this method to consider what happens to the late time slope under different distributions of t_{cb}. This is discussed in Chapter 7.

5.3.3 Method of moments applied to DP model

5.3.3.1 Moments of the BGF

It is useful to consider both the moments for the impulse response of the block (Table 5.7) as well as for the BTC (Table 5.8). Differentials of the BGF are hereafter written as $\hat{B}^{(n)}(0,t_{cb}) = \frac{d^n}{ds^n} B\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$.

Table 5.7: Temporal moments of the BGF, where $t_{cb} = b^2/D_a$ and $\hat{B}^{(n)}(0,t_{cb}) = \frac{d^n}{ds^n} B\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$.

	$B\left(\sqrt{st_{cb}}\right)$	$B\left(\sqrt{st_{cb}}\right)\Big\|_{s=0}$	$\mu_1' = -\frac{d}{ds} B\left(\sqrt{st_{cb}}\right)\Big\|_{s=0} = -\hat{B}'(0,t_{cb})$	$\mu_2' = -\frac{d^2}{ds^2} B\left(\sqrt{st_{cb}}\right)\Big\|_{s=0} = -\hat{B}''(0,t_{cb})$	$\mu_3' = -\frac{d^3}{ds^3} B\left(\sqrt{st_{cb}}\right)\Big\|_{s=0} = -\hat{B}'''(0,t_{cb})$	μ_2	μ_3
Slab $b = h$	$\tanh(\sqrt{t_{cb}s})/\sqrt{t_{cb}s}$	1	$\frac{t_{cb}}{3}$	$\frac{4}{15}t_{cb}^2$	$\frac{34}{105}t_{cb}^3$	$\frac{7}{45}t_{cb}^2$	$\frac{124}{945}t_{cb}^3$
Cylinder $b = r/2$	$\frac{I_1(2\sqrt{t_{cb}s})}{\sqrt{t_{cb}s}I_o\left(2\sqrt{t_{cb}s}\right)}$	1	$\frac{t_{cb}}{8}$	$\frac{1}{24}t_{cb}^2$	$\frac{11}{512}t_{cb}^3$	$\frac{5}{192}t_{cb}^2$	$\frac{5}{512}t_{cb}^3$
Sphere $b = r/3$	$\coth\left(3\sqrt{t_{cb}s}\right)/\sqrt{t_{cb}s} - 1/(3t_{cb}s)$	1	$\frac{t_{cb}}{15}$	$\frac{4}{315}t_{cb}^2$	$\frac{2}{525}t_{cb}^3$	$\frac{13}{1575}t_{cb}^2$	$\frac{44}{23625}t_{cb}^3$
Infinite matrix	$1/\sqrt{st_{cb}}$	∞	∞	∞	∞	∞	∞
MIM	$\frac{\kappa}{\theta_{im}s + \kappa}$	1	$\frac{\theta_{im}}{\kappa} = t_M$	$\frac{2\theta_{im}^2}{\kappa^2}$	$6\frac{\theta_{im}^3}{\kappa^3}$	$\frac{\theta_{im}^2}{\kappa^2}$	$2\frac{\theta_{im}^3}{\kappa^3}$

5.3.3.2 Moments of the DP model BTC

Table 5.8: Temporal moments of the BTC. *

Geometry	$\mathrm{B}\left(\sqrt{st_{cb}}\right)$	$\mu_0 = \mathrm{B}\left(\sqrt{st_{cb}}\right)\Big\|_{s=0}$	$\mu_1' = t_a(1+\sigma)$	$\mu_2 = -2t_a\sigma\hat{\mathrm{B}}'(0,t_{cb})$	$\mu_3 = 3t_a\sigma\hat{\mathrm{B}}''(0,t_{cb})$
Slab $D_C = b$ t_{cb0}	$\tanh(\sqrt{t_{cb}s})/\sqrt{st_{cb}}$	m_0/V_m	$t_a(1+\sigma)$	$2t_a\dfrac{\sigma t_{cb}}{3}$	$-4t_a\sigma\dfrac{t_{cb}^2}{5}$
Cylinder $D_C = r/2$ $t_{cb} = t_{cb0}/4$	$\dfrac{\mathrm{I}_1(2\sqrt{st_{cb}})}{\sqrt{st_{cb}}\,\mathrm{I}_o(2\sqrt{st_{cb}})}$	m_0/V_m	$t_a(1+\sigma)$	$t_a\dfrac{\sigma t_{cb}}{4}$	$-t_a\sigma\dfrac{t_{cb}^2}{8}$
Sphere $D_C = r/3$ $t_{cb} = t_{cb0}/9$	$\coth\left(3\sqrt{st_{cb}}\right)\Big/\sqrt{st_{cb}} - 1/(3st_{cb})$	m_0/V_m	$t_a(1+\sigma)$	$2t_a\dfrac{\sigma t_{cb}}{15}$	$-4t_a\sigma\dfrac{t_{cb}^2}{105}$
Infinite matrix	$1/\sqrt{st_{cb}}$	∞	$t_a(1+\sigma)$	∞	∞
MIM	$\dfrac{\kappa}{\theta_{im}s+\kappa}$	m_0/V_m	$t_a(1+\sigma)$	$2t_a\dfrac{\sigma\theta_{im}}{\kappa}$	$-6t_a\sigma\dfrac{\theta_{im}^2}{\kappa^2}$

* **where,** $\hat{\mathrm{B}}'(0,t_{cb}) = \dfrac{d}{ds}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$ **[T],** $\mathrm{B}''(0,t_{cb}) = \dfrac{d^2}{ds^2}\hat{\mathrm{B}}\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$ **[T²] or alternatively**

$\hat{\mathrm{B}}'(0) = \dfrac{d}{dS}\mathrm{B}\left(\sqrt{S}\right)\Big|_{S=0}$ **and** $\mathrm{B}''(0) = \dfrac{d^2}{dS^2}\hat{\mathrm{B}}\left(\sqrt{S}\right)\Big|_{S=0}$ **and so on.**

The moments can now be evaluated by substitution into the following expressions:

$\mu_1' = t_a(1+\sigma)$ [T]

$\mu_2 = -2t_a\sigma\hat{\mathrm{B}}'(0,t_{cb}) = -2t_at_{cb}\sigma\hat{\mathrm{B}}'(0)$ [T²]

$\mu_3 = 3t_a\sigma\hat{\mathrm{B}}''(0,t_{cb}) = 3t_at_{cb}^2\sigma\hat{\mathrm{B}}''(0)$ [T³],

$\mu_4 = 12t_a^2\sigma^2\hat{\mathrm{B}}'(0,t_{cb})^2 - 4t_a\sigma\hat{\mathrm{B}}'''(0,t_{cb}) = 12t_a^2t_{cb}^2\sigma^2\hat{\mathrm{B}}'(0)^2 - 4t_at_{cb}^3\sigma\hat{\mathrm{B}}'''(0)$ [T⁴].

So the first moments are independent of B, the second (centred) moments are governed by $\hat{\mathrm{B}}'(0)$, the third (centred) moments are governed by $\hat{\mathrm{B}}''(0)$ and the fourth (centred) moments are governed by $\hat{\mathrm{B}}'(0)$ and $\hat{\mathrm{B}}'''(0)$.

The fact that the first BTC moment (the mean arrival time) is independent of the BGF (and t_{cb}) is a very useful and general result. This is generalized by the 'equal time theorem' (Redner (2001) states this as, "the mean first-passage time across a network equals the system volume divided by the total external flow").

The fact that the second BTC moment is only a function of the first BGF moment is also very useful. This raises the possibility of comparing different BGFs so as to enable the matching of the spread of the BTC.

Infinite blocks give infinite moments (except for the first moment). This makes matching by the method of moments impossible for infinite blocks.[37]

5.3.4 Monte Carlo used for examining the DP model by approach 1

In a similar vein to the treatment for the AD model in Section 5.2, the DP model is now compared to synthetic data generated by the same DP model under different process conditions (approach 1 of Chapter 4). The application of the MC method proved to be more problematic for the DP than the AD model for a number of reasons.

It is not practical to present a comprehensive range of parameter combinations. Instead a few arbitrary scenarios are used to illustrate the typical inferences that might be made using this method. The automatic fitting proved to be more difficult than for an AD system, particularly at very large or very small t_{cb}/t_a ratios. The MC examination is split into sections examining the effects of noise, target parameters and parameterisation. In 'Approach 1', MC method is used to repeatedly fit models to synthetic data which are created by adding noise to a model-generated BTC by randomly selecting from a normal distribution, $\mathrm{N}\left(0,\sigma_N^2\right)$. The distribution of parameters estimated this way is presented in a 2D scatter plot and CRs are calculated.

[37] It is an interesting but unanswered question as to what would define a wider class of BGFs that would give infinite moments and whether any general statements can be made relating such BGFs.

5.3.4.1 Effect of noise

Whilst even for a noise of magnitude $\sigma_N = 0.01(m_0/V)$ the AD model produced elliptical MC scatter-plots for all parameter combinations trialled, this was not the case for the DP model. Only when the noise was substantially reduced were elliptical confidence regions (CRs) generated. Therefore any results based purely on the Hessian (such as the CIs generated by PEST and CXTFIT) should be viewed cautiously. Any trends established from the CI produced this way should ideally be checked against a MC plot. This suggests that the CI given by automatic-fitting packages should be used with care.

Compare a series of MC scatter-plots generated for the following (arbitrary) conditions, with absolute noise steadily increasing: z=50; $t_{cb} = 1$; $\sigma = 1$; $\theta = 0.2$; $q = 1$. First, examining the SSEs near to the 'target' point confirms that the error space is elliptical very close to the 'target' point (Figure 5.11) as would be expected for any model. However, the error ellipse is curved further away from this (Figure 5.12).

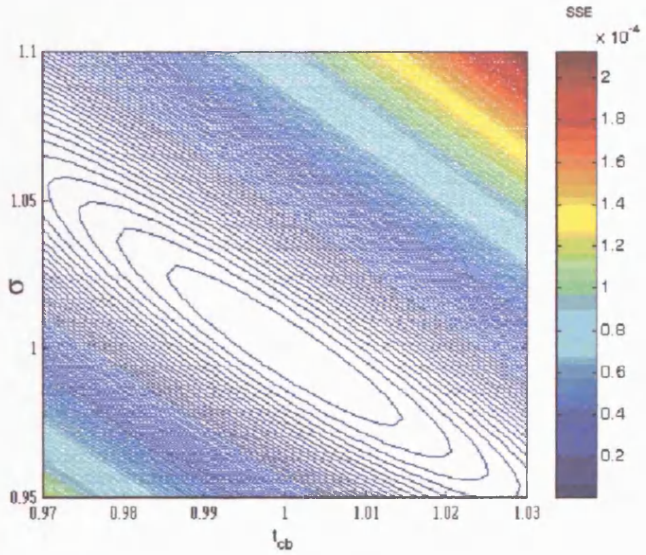


Figure 5.11: SSE of $C/(m_0/V_m)$ near target for $z = 50$, $t_{cb} = 1$, $\sigma = 1$, $\theta = 0.2$, $q = 1$.

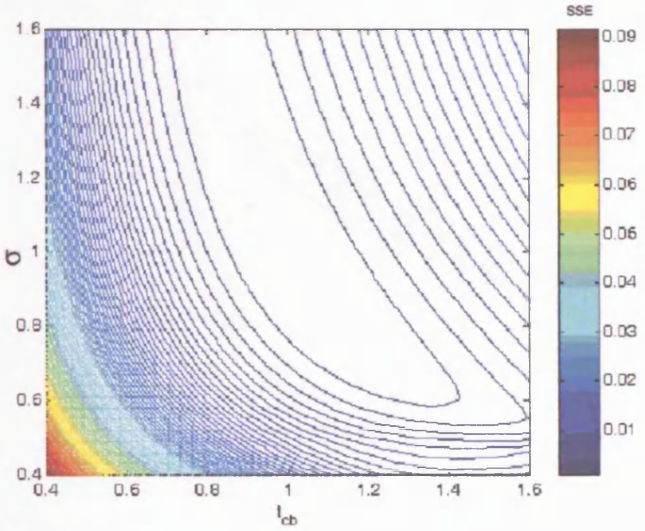


Figure 5.12: SSE of $c/(m_0/V_m)$ further from 'target' for $z = 50$, $t_{cb} = 1$, $\sigma = 1$, $\theta = 0.2$, $q = 1$.

The difficulty with examining error space is in knowing how far from the 'target' the CR of interest lies (most probably the 95% CR). It therefore requires the MC analysis to produce scatter plots in order to define how far from the target is of interest. The MC analysis indeed showed that an elliptical assumption is only valid for relatively low levels of noise (Figures 5.13, 5.14 and 5.15).

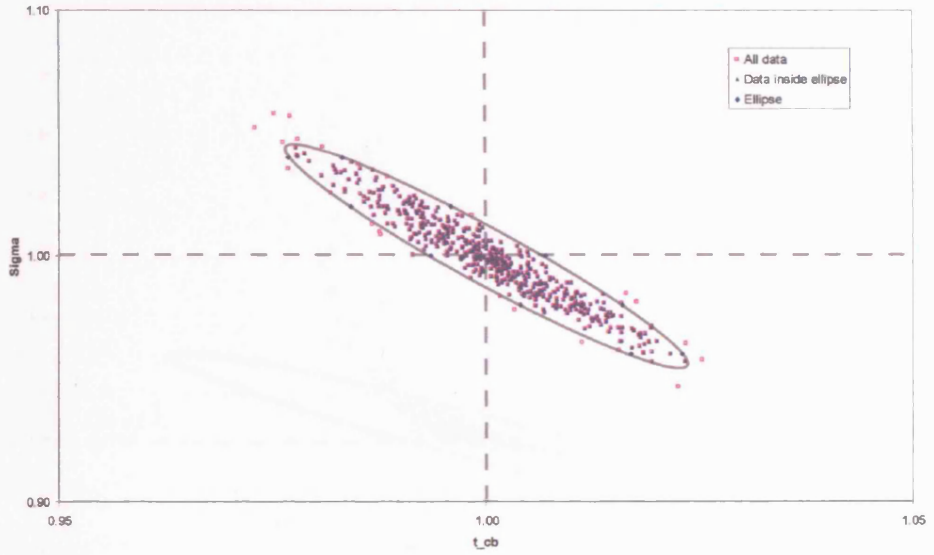


Figure 5.13: MC for $\sigma_N = 0.001(m_0/V)$, where 95% points were inside the ellipse, for 500 realisations.

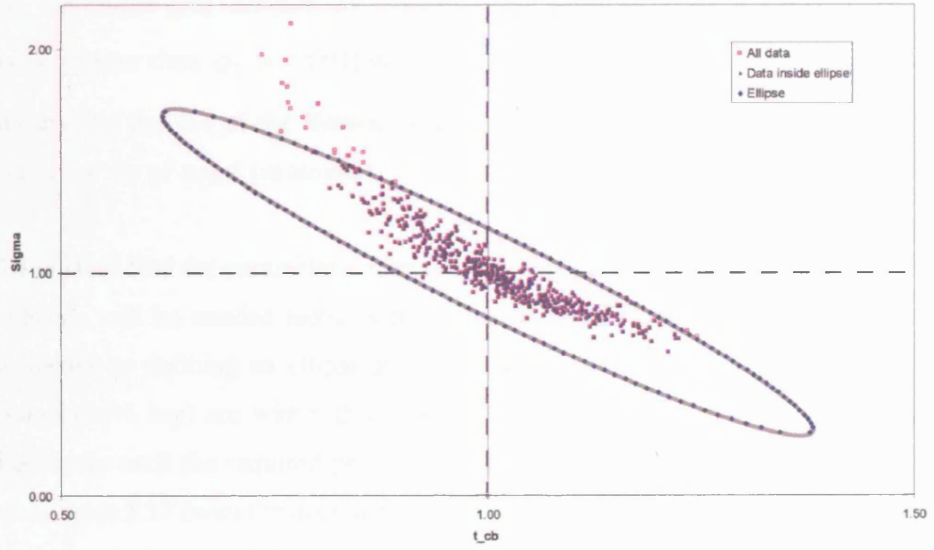


Figure 5.14: MC for $\sigma_N = 0.01(m_0/V)$, where 87.6% points were inside the ellipse, for 500 realisations.

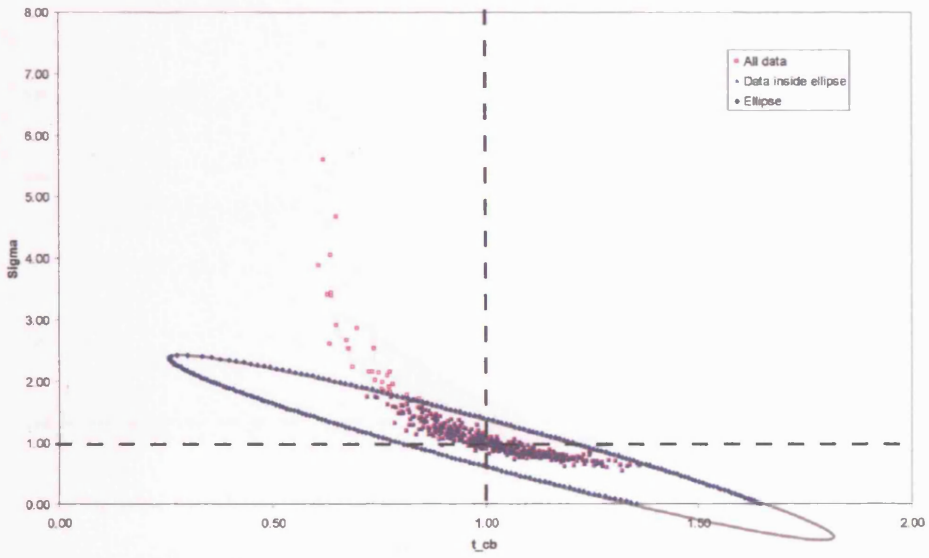


Figure 5.15: MC for $\sigma_N = 0.015(m_0/V)$, where 81.6% points were inside the ellipse, for 500 realisations.

So, the ellipse gets increasingly 'curved' away from the target as the noise level increases. For noise greater than $\sigma_N = 0.001(m_0/V)$, the elliptical assumption is no longer reasonable. This means that the use of the Hessian to estimate CIs or elliptically-shaped CRs is limited for this particular set of target parameters.

To reliably find the parameter uncertainties for systems with noise above this threshold pure MC analysis will be needed along with a new assumed shape of CR. A better approximation is achieved by defining an ellipse on log-log axes. In order to ensure the required percentage of points (95% say) are within this new CR, the size of this log-log ellipse needs to be adjusted iteratively until the required percentage are within the CR. Examples are given in Figure 5.16 and Figure 5.17 (with the axes kept unlogged to demonstrate the change in shape of the CR). It is clear that the log-log ellipse describes the curved CR better, but still imperfectly.

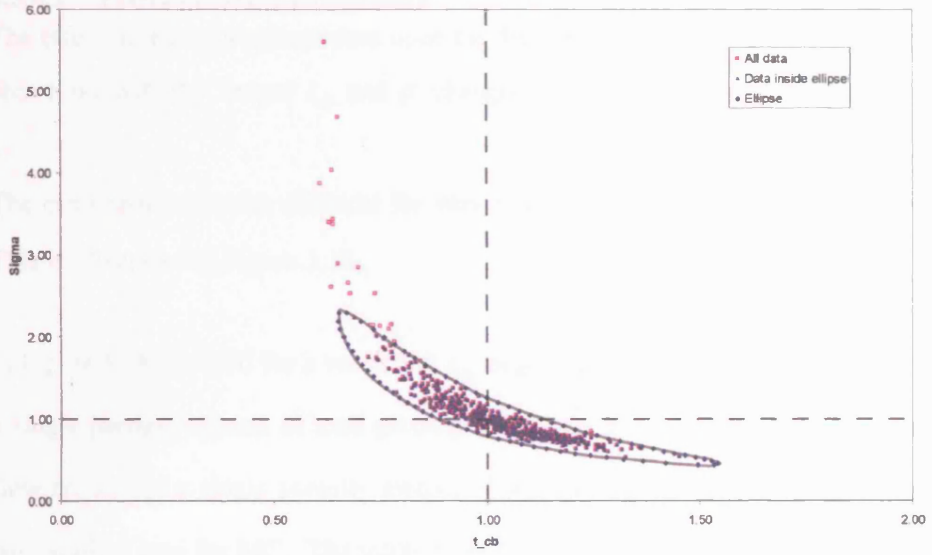


Figure 5.16: MC for $\sigma_N = 0.015(m_0/V)$ with log-log

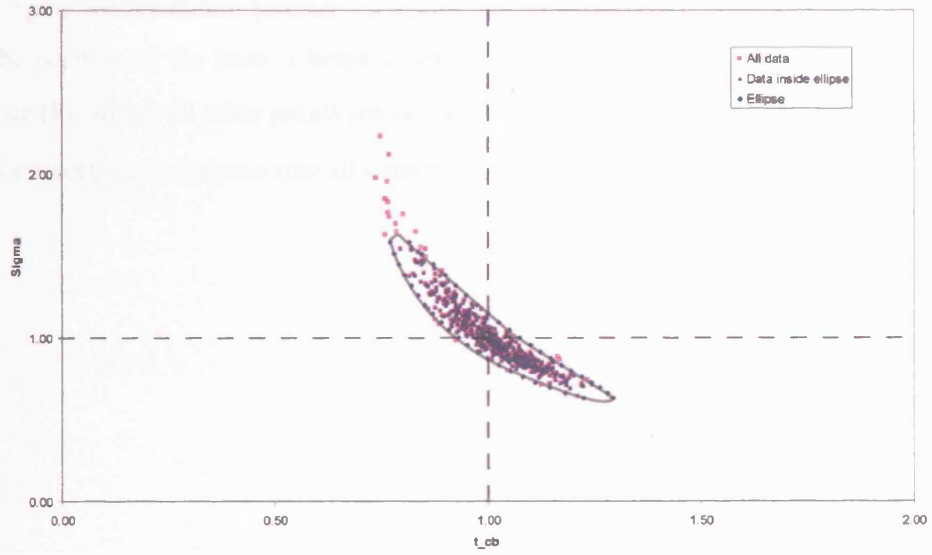


Figure 5.17: MC for $\sigma_N = 0.01(m_0/V)$ with log-log axes

The log-log ellipse therefore improves the situation after the normal ellipse is inadequate, but for higher values of error it is itself a poor replication of the underlying distribution.

5.3.4.2 Effect of target parameters

The effect of the target parameters upon the distribution is examined, for the similar conditions to above, but with the 'target' t_{cb} and σ changed for a top-hat pulse of duration $T_p = 10$ days.

The error space remains elliptical for very large t_{cb}, but becomes elongated for very short t_{cb}. This is illustrated in Figure 5.18.

In Figure 5.18 the BTC for a very short t_{cb} begins to look very much like a piston-flow arrival for a single porosity system of total porosity θ and the BTC for very long t_{cb} looks like a piston-flow arrival for a single porosity system of porosity θ_m. Such pure piston-flow is potentially a pathological case for MC. The solute front from C_0 to 0 or from 0 to C_0 becomes sharper than the gaps between datapoints (Δt). The fitting algorithm may not adjust parameters sufficiently to move the position of the front to the next time node and the parameter may appear to be completely insensitive.

Upping the resolution (shorter Δt) does not overcome this. This is probably because although the position of the front is better constrained the number of fitting points increases at the same rate (for which all other points are either C_0 or 0). Least-squares is therefore not a good method for recording the precise time of a sharp front.

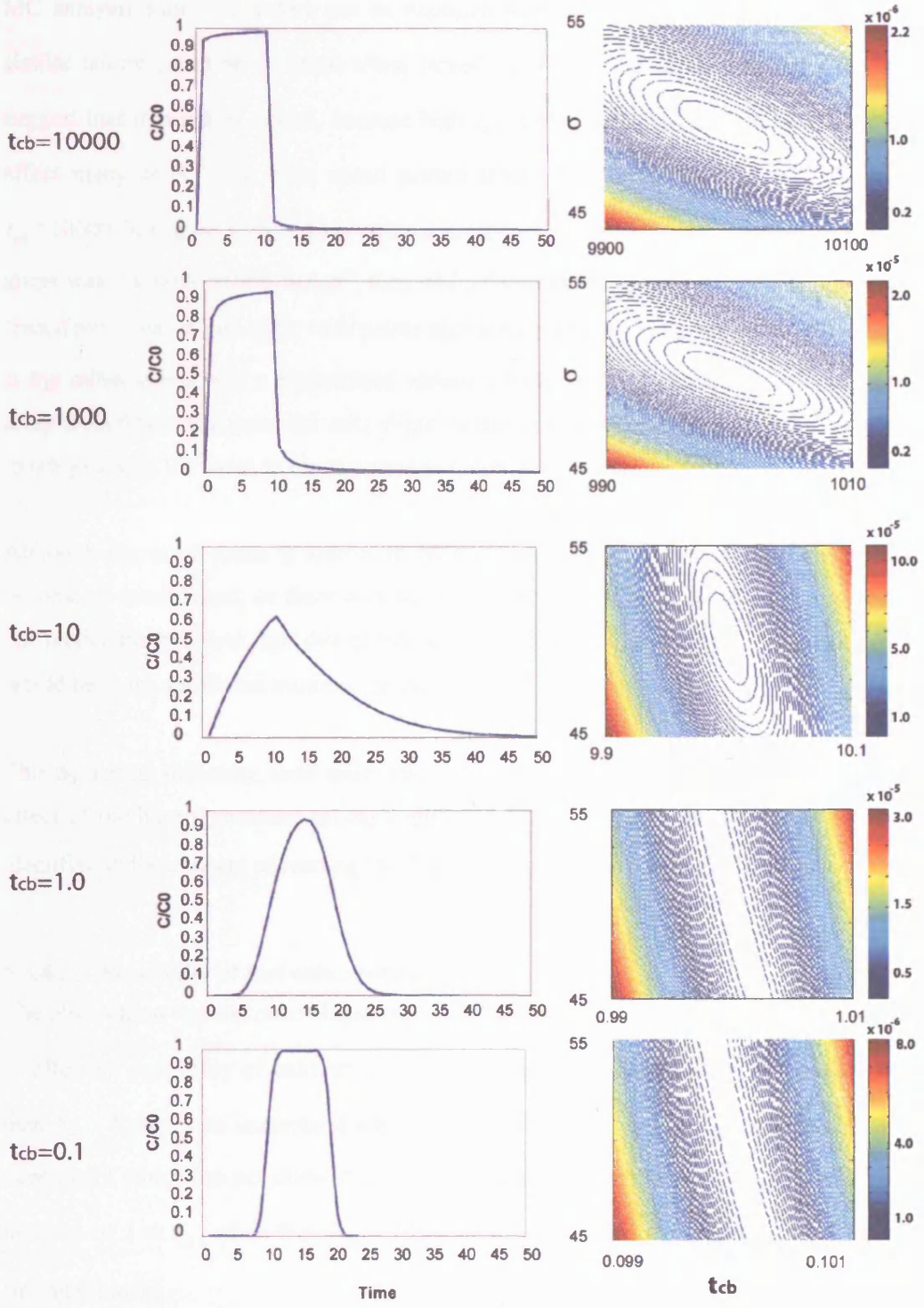


Figure 5.18: Comparison of error space for $z = 50$, $\sigma = 50$, $\theta = 0.2$, $q = 1$ and variable t_{cb}.

MC analysis should therefore not be expected to work very well when 'target' $t_{cb} \ll t_a$. A similar failure could be expected when 'target' $t_{cb} \gg t_a$. However, examining the BTC would suggest that it is not as severe, because high t_{cb} cases have a tail, for which changes in the t_{cb} affect many cells. The error space plotted around the 'target' for $t_{cb} \gg t_a$ is elliptical (e.g t_{cb}=10000 in Figure 5.18). This is in contrast to $t_{cb} \ll t_a$. In the latter instance, if the initial guess was the same as the 'target', then odd patterns emerged on the scatter plots, for example a 'cruciform' pattern emerged, with points appearing primarily in the direction of either t_{cb} or σ. If the initial guess was a randomised variation from the 'target', then there was no movement away from this initial guess (or only slight variations from this). This led to the apparent solution looking exactly the same as the guesses (an elliptical bivariate normal).

Although the error space is known to be elliptical it may be that the minimising routine has become ill-conditioned, or there may be an accuracy problem. Calibrating in logged space did not rectify the problem, nor did adjusting the minimiser's parameters. One possible next step would be to try a different minimising algorithm.

This numerical difficulty, until overcome, removes the ability for systematic investigation of the effect of the base parameters on the width of the CI. However, no theoretical barrier has been identified to the concept of running much shorter tests than the t_{cb} that is sought.

5.3.4.3 Possibility of fast column-tests

The observation that the error ellipse remains elliptical when $t \ll t_{cb}$ is remarkable. This appears to offer the possibility of calibrating t_{cb} using data from experiments which are much shorter than t_{cb}. In order to understand why it might be possible to establish the block time with an experiment which has not allowed solute to diffuse to the centre of the blocks, first consider the instance of $t \ll t_{cb}$, when the BGF is approximated as $\mathrm{B}\left(\sqrt{st_{cb}}\right) \sim 1/\sqrt{st_{cb}}$ (i.e. an effectively 'infinite' block).

Compare two ways to express the DP response, either with t_a fixed:

$$\bar{C}_m = (m_0 / V_m)\exp\left[-st_a\left(1+\frac{\sigma}{\sqrt{st_{cb}}}\right)\right], \text{ or since } t_{cf} = \frac{t_{cb}}{\sigma^2}$$
$$\bar{C}_m = (m_0 / V_m)\exp\left[-st_a\left(1+\frac{1}{\sqrt{st_{cf}}}\right)\right] \quad (5.36)$$

or, with q and θ fixed:

$$\bar{C}_m = (m_0 / V_m)\exp\left[-\frac{z\theta}{q(1+\sigma)}s\left(1+\frac{1}{\sqrt{st_{cf}}}\right)\right] \quad (5.37)$$

For fixed t_a in Equation (5.36), calibration will find the ratio $\frac{\sigma}{\sqrt{t_{cb}}} = \frac{1}{\sqrt{t_{cf}}}$. In other words, only t_{cf} is found and a line of non-uniqueness is expected in t_{cb} vs. σ space.

For t_a re-expressed in terms of σ (Equation (5.37)) calibration will find σ and t_{cf}, thereby also constraining t_{cb}, despite the fact that $t \ll t_{cb}$. So, even if the diffusion 'sees' an infinite matrix, both σ and t_{cb} are still recovered. This apparently contradicts reasoning given by Ward et al. (1998) and Haggerty (1999), that t_{cb} cannot be revealed when $t \ll t_{cb}$ or when $t \gg t_{cb}$ and that a test should be of the order of t_{cb} in order to reveal t_{cb}. However, this conventional wisdom refers to field experiments, where the domain boundaries θ and $Q = qA$ are not well-constrained, yet t_a is a convenient and meaningful parameter to work with. However, column tests and certain other 1D systems do allow reparameterisation of t_a.

For experiments where $t \ll t_{cb}$ is no longer true, calibration of the equation describing exchange with a finite block is expected to recover σ and t_{cb}. The finite block transport equation is reiterated here as:

$$\bar{C}_f = (m_0 / V_m)\exp\left[-st_a\left(1+\sigma \mathrm{B}\sqrt{st_{cb}}\right)\right] \quad (5.38)$$

Because expressing t_a in terms of σ gives greater constraint to σ for the infinite block situation, it would be expected to continue to do this for the finite block case as well. Examination of the error space in Figure 5.19 confirms this anticipated result.

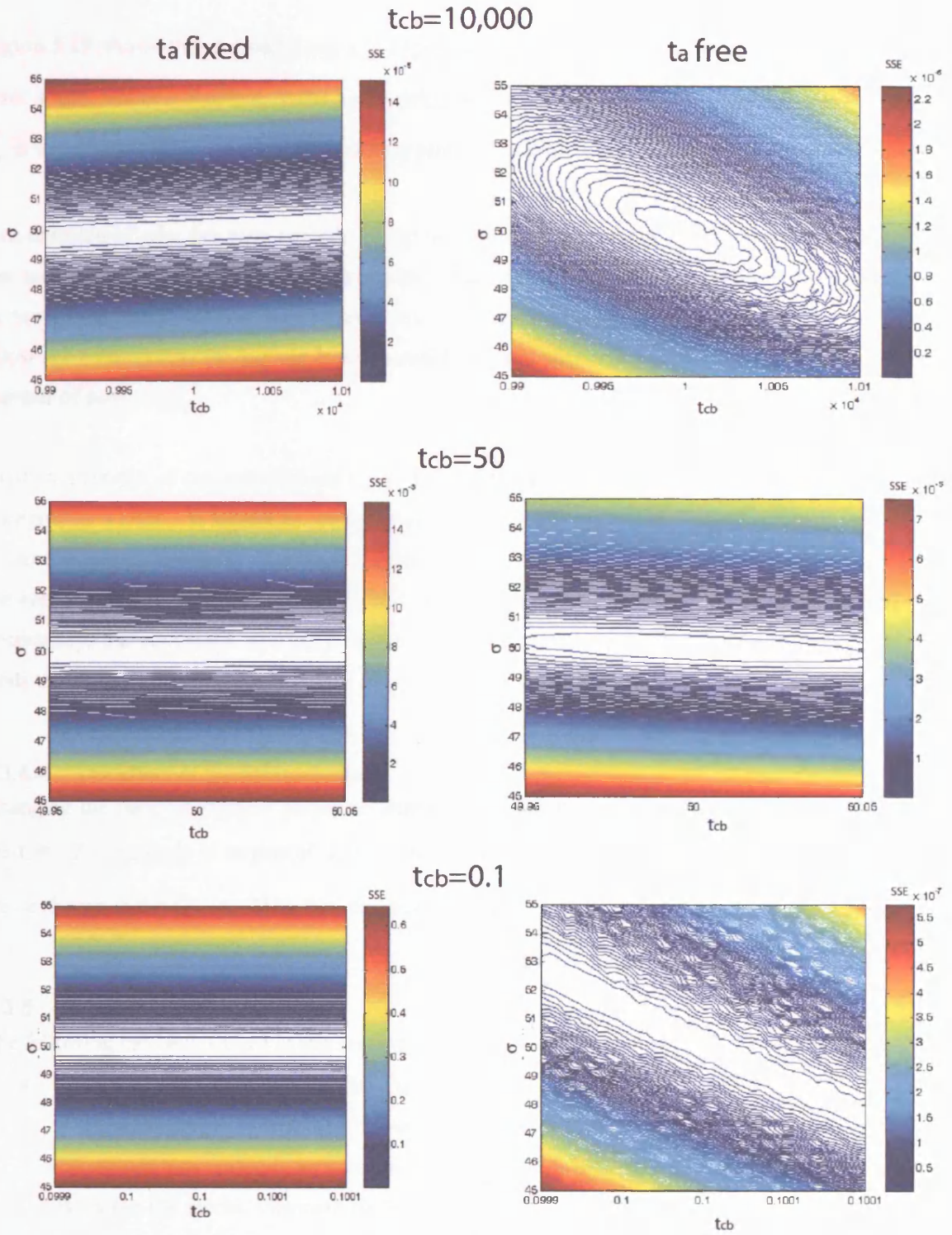


Figure 5.19: Contours of SSR for pulse injection with t_a fixed and free compared (L =50, $\theta = 0.2$, $q = 1$ and T_p =10. The 'base' $\sigma = 50$ and base t_a is therefore 0.2).

Figure 5.19 shows that for very long t_{cb}, expressing t_a as a function of σ yields an elliptical error space, whereas fixing t_a gives a non-uniqueness. This improvement continues down until t_a is of the order of t_{cb}. For $t_{cb} \ll t_a$ neither parameterisation gives an elliptical error space.

One example of why fast tests could be useful would be that of batches of laboratory tests where the transport model itself is not being tested but instead how the parameters change with systematic changes in process conditions (e.g. compressive stress) is tested. Fast tests would allow more datapoints for a given experimental time, or a shorter experimental time for a given number of datapoints.

Another example of the potential utility of fast tests would be for non-stationary systems like undegraded wastes. It would be useful to provide a quick 'snapshot' of transport parameters before the system changes. The main proviso would be not to let the flow rates go so high that the system assumes a completely new flow geometry or flow law. Also in the instance of process uncertainty, the advantage of a long experiment is that it tests both the conceptual model *and* finds the parameters.

5.3.4.4 The effect of parameterisation

Changing the parameterisation makes no difference to the estimated percentage of points within the CR. For example if targets of t_{cb}=10 and σ=1 were replaced by t_{cb}=10 and t_{cf}=10, the major change in the CR would be that the target for the y-axis was changed from σ=1 to t_{cf}=10.

5.3.5 Conclusions

The following has been shown in this section:

- Similar amounts of noise that for the AD model fitted to synthetic data generated by an AD model gave elliptical distributions of parameters, but gave curved ellipses when the DP system was tested in an identical way. The CR is elliptical only for very small error is for the DP model. Standard models such as PEST and CXTFIT which compute the CI based on the covariance matrix should therefore be used with some caution.
- The 'target' point affects the shape of the error space. In particular a very short t_{cb} generates a piston-flow which presents a 'pathological' case to the MC method.

It is worth considering what should be done for future (real) calibrations of DP models, where it is important to establish the CIs for different parameters. An assessment can be made of the level of noise (for example by analysis of the best-fit residuals). If the noise is very small, then the chances of inaccurate CR determination are reduced, otherwise the estimate of noise could then be used to conduct an MC analysis in the same way as has been performed here (except that random noise would be repeatedly added to the best-fit model). If the error in the proximity of the contour describing the 95% CR is not elliptical, then there are a range of options that could be adopted. These are:

- Using the log-log method to define a better CR (which is checked visually).
- Defining a new CR (e.g. selecting an appropriate shape which is sized in order to contain 95% of points).
- Repeating the experiment having identified and reduced contributions to noise.
- If the CR was being established in order to allow future prediction, taking in to account parameter uncertainty, then consideration may be needed of methods which do not rely on the CR alone. For example, one way of generating a distribution of predictions would be to take a simple bootstrap from the calibrated MC distribution in a 'forward' MC model.

5.4 *Non-linear sorption*

The coupling of non-linear sorption with other processes is the least readily amenable to analytical analysis out of all the processes so far considered. The non-linearity diminishes the utility of the many techniques that rely on linearity (especially the use of superposition, LTs and their application to the method of moments). Investigation of non-linear coupled models has therefore tended to be a numerically-based exercise, although some special cases amenable to analytical treatment are examined here.

Kinetic sorption models, when equilibrium is not instantaneous, are potentially mathematically equivalent to DP(MIM) exchange models and are therefore not considered independently. This equivalence is discussed in detail by Haggerty and Gorelick (1998).

The coupling of non-linear sorption with other processes remains a challenge (Sardin et al., 1991), although there has now been work in this area. For example, Brusseau (1995) investigated non-linear sorption in conjunction with degradation.

5.4.1 General non-linear effect

A general sorption isotherm is defined as $C^* = \mathrm{f}(C)$, where C^* is the sorbed mass per unit mass of dry solid. The chain rule gives $\frac{\partial C^*}{\partial t} = \frac{\partial C^*}{\partial C}\frac{\partial C}{\partial t} = \mathrm{f}'(C)\frac{\partial C}{\partial t}$.

Consider advection and sorption in the absence of mechanical dispersion or DP exchange,

$$\frac{\partial C}{\partial t} = -V\frac{\partial C}{\partial z} - \frac{B_d}{\theta}\frac{\partial C^*}{\partial t} \tag{5.39}$$

where B_d is bulk density and θ is the total porosity.

Substituting for C^* and rearranging gives,

$$\left(1+\frac{B_d}{\theta}\mathrm{f}'(C)\right)\frac{\partial C}{\partial t} = -V\frac{\partial C}{\partial z} \tag{5.40}$$

For a given constant C the characteristic wave velocity, V_C, is therefore:

$$V_C = \left.\frac{dz}{dt}\right|_C = \frac{V}{1+\frac{B_d}{\theta}\mathrm{f}'(C)} \tag{5.41}$$

Clearly the velocity will vary with $C(z,t)$, so at any instance the different parts of a concentration profile will advance at different speeds, causing the front to either sharpen or spread. This will depend on $\mathrm{f}'(C)$ and the slope of the front (see Table 5.9 and Figure 5.20 and Figure 5.21).

Table 5.9: The effect of isotherm curvature and slope of the concentration front ($C_1(z)$) on the subsequent sharpening or spreading of the front.

Initial solute front gradient	Isotherm curvature	Isotherm shape	Effect on solute fronts
$\frac{d}{dz}C_1(z) > 0$	$f''(C) > 0$	Convex	Back catches up the front: ***sharpening***
	$f''(C) < 0$	Concave	Back left behind by the front: ***spreading***
$\frac{d}{dz}C_1(z) < 0$	$f''(C) > 0$	Convex	Back left behind by the front: ***spreading***
	$f''(C) < 0$	Concave	Back catches up the front: ***sharpening***

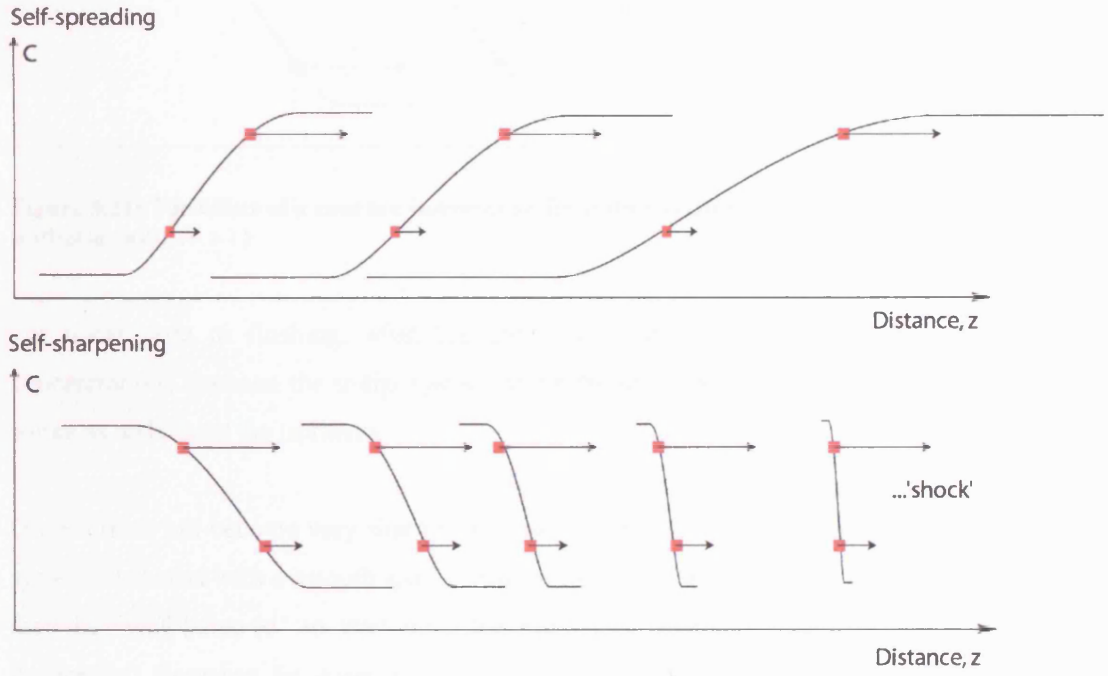


Figure 5.20: The effect of a convex isotherm on front dispersion/sharpening (e.g. Freundlich isotherm, with $N < 1$)

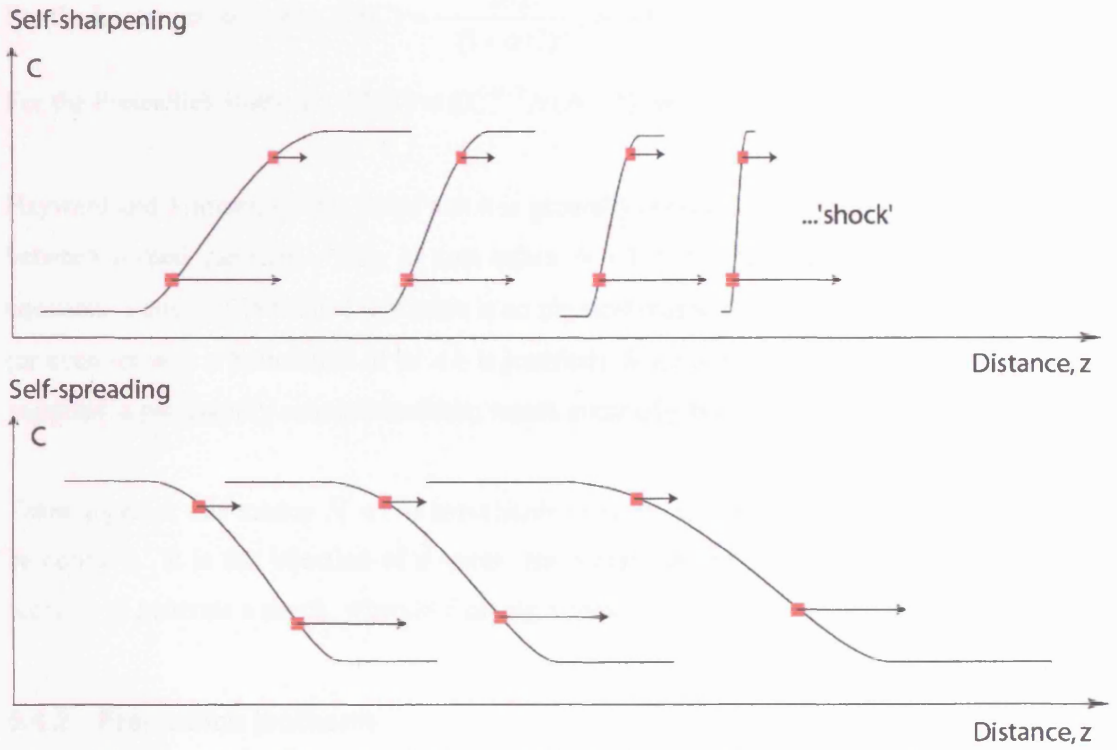


Figure 5.21: The effect of a concave isotherm on front dispersion/sharpening (e.g. Freundlich isotherm, with $N > 1$)

For tracer tests or flushing, what happens to a front will therefore depend on the relative concentrations between the initial system concentration ('background') and that of the invading solute as well as on the isotherm.

Once a front has become very sharp it is called a 'shock'. In the instance of a sharpening front, systems that start with a smooth spatial distribution of tracer (such as the example in Figure 5.20) may be 'well-behaved' to start with but may then potentially develop shocks as transport progresses. However, for systems which initially have solute within them and are then flushed or are initially clear and suddenly have a tracer injected there is the potential for a shock to be generated at the start.

Specifically, conditions that are likely to generate shocks are:

- Injection of flush into solute-filled system and a convex isotherm ($f''(C) > 0$).
- Injection of solute into clean system and a concave isotherm ($f''(C) < 0$).

For the Langmuir isotherm, $f''(C) = -\frac{2\alpha\beta}{(1+\alpha C)^3}$, which is always <0.

For the Freundlich isotherm, $f''(C) = KC^{N-2}N(N-1)$, which is <0 if $N<1$ or >0 if $N>1$.

Hayward and Trapnell (1964) noted that it is generally observed that there is increasing repulsion between sorbed particles. This in turn infers $N<1$ from their derivation of the Freundlich equation. Fetter (1999) stated that there is no physical reason for $N<1$, but without explanation (or even for why a Freundlich of $N<1$ is justified). Since a maximum sorption capacity is to be expected, a persistently concave isotherm would eventually become unphysical.

Taken together this means $N<1$ is most likely to occur, so both types of isotherms are likely to be concave. It is the injection of a tracer into a clean system that is therefore the most likely scenario to generate a shock, whereas flushing is most likely to give spreading.

5.4.2 Freundlich isotherm

The specific case of a Freundlich isotherm is examined in more detail here. The reason it is selected here is that it is simple, yet often effective at modelling real data. Whether it remains relevant to very low concentrations will depend on the particular dataset. The analysis that follows is intended for use with concentration ranges in which the isotherm has been established or simply for testing out what is predicted to occur if a Freundlich isotherm *is assumed.*

Limits can be established for the velocities of extreme concentration fractions:

Substituting $f'(C) = KNC^{N-1}$ into Equation (5.41) yields:

$$V_C = \frac{V}{1+\lambda NC^{N-1}}, \text{ where } \lambda = \frac{B_d K}{\theta} \tag{5.42}$$

So, for $0<N<1$: $\lim_{C\to 0}\left.\frac{dz}{dt}\right|_C = 0$ and $\lim_{C\to\infty}\left.\frac{dz}{dt}\right|_C = V$.

Conversely, for $N>1$: $\lim_{C\to 0}\left.\frac{dz}{dt}\right|_C = V$ and $\lim_{C\to\infty}\left.\frac{dz}{dt}\right|_C = 0$.

This maximum range in velocities limits the rate that dispersion or sharpening occurs.

5.4.3 Shock interface

For $N<1$ and when the initial concentration exceeds the injected concentration ($C_I > C_0$) an undispersed shock or 'step' of tracer will invade the system. Any slight dispersion at the front will be counteracted by the fact that any upstream-dispersed particles will have lower concentration and therefore lower velocity than the centre of the front and will be caught up. Note that shocks also appear in aerodynamic systems with compressible flow, for hydraulic jumps and for Buckley-Leverett multiphase flow.

The shock velocity is given by $V_S = \dfrac{V}{1+\dfrac{\Delta f}{\Delta C}}$, where $\Delta f = \dfrac{B_d K}{\theta}\left(C_I^N - C_0^N\right)$ and $\Delta C = C_I - C_0$ (Wu et al., 1997; van der Zee, 1990). It is straightforward to demonstrate that this velocity satisfies mass balance. In the instance of a continuous input and a shock advancing into the medium, V_S can be confirmed by cancelling and rearranging the mass balance equation that arises from the statement that mass injected equals change in mass in the system. So:

$$AV\theta C_I \Delta t - AV\theta C_0 \Delta t = AV_S \Delta t\left(B_d K C_I^N + \theta C_I\right) - AV_S \Delta t\left(B_d K C_0^N + \theta C_0\right) \qquad (5.43)$$

A formal proof that no other valid mathematical solution exists has not been found. In particular it would be desirable to prove whether or not $C/C_0 \leq 1$ for all z, t.

In reality dispersive effects tend to work against the build-up of sharp gradients (the $D\dfrac{\partial^2 C}{\partial z^2}$ term in the ADE). Eventually the shock forms a stable smoother profile which represents the balance of these two counteracting effects. Van der Zee (1990) presented analytical solutions for the shape of the shock (somewhat complicated by the addition of a second non-equilibrium sorption process).

5.4.4 Dispersed interface

A corollary to the shock interface for the exact opposite set of circumstances is a spreading boundary.[38] For example, when $N<1$ and a step input of solute free solute is injected, for all

[38] It is straightforward to derive a wave speed and solution for the spreading wave for a Langmuir isotherm in a similar manner as for the Freundlich isotherm.

$t > 0$ the concentration of a tiny element of column Δz immediately adjacent to the inlet will become equal to 0. This 'particle' and all subsequent particles entering the system will move at the minimum speed $\left.\frac{dz}{dt}\right|_{C_0} = 0$. Any slight dispersion at the front will result in lower concentration particles at the back which will then move at slower velocity, thereby being left behind by the main plume. The flushing front becomes increasingly dispersed.

In the absence of mechanical dispersion, a simple analytical solution is available. Consider a system with a Freundlich isotherm, a step input and a spreading front described by:

$$\left(1+\lambda N C^{N-1}\right)\frac{\partial C}{\partial t} + V\frac{\partial C}{\partial z} = 0 \tag{5.44}$$

The characteristic velocity is therefore:

$$V_C = \left.\frac{dz}{dt}\right|_C = \frac{V}{\left(1+\lambda N C^{N-1}\right)} \tag{5.45}$$

For $N > 1$ and a step input boundary condition $C(z,0) = 0$, $C(0,t) = 1$ the velocity of $C = 0$ is $V_{C_0} = V$ and the velocity of $C = 1$ is $V_{C_1} = \frac{V}{(1+\lambda N)}$. A solution to Equation (5.44) for these conditions is given by:[39]

$$\begin{aligned} C(z,t) &= \left(\frac{Vt/z - 1}{\lambda N}\right)^{1/(N-1)} \quad \text{for } t < (1+\lambda N C_0^{N-1})z/V, \\ C(z,t) &= 0 \text{ for } t < z/V \text{ and} \\ C(z,t) &= C_0 \text{ for } t > (1+\lambda N C_0^{N-1})z/V \end{aligned} \tag{5.46}$$

The limits to t for which the different parts of the solution are valid can be explained by considering the velocities of different concentrations. C_0 moves the slowest and arrives at time $t = \left(1+\lambda N C_0^{N-1}\right)z/V$ and $C = 0$ moves the fastest and arrives at time $t = z/V$.

[39] This has also been checked in Derive® as a valid mathematical solution to the transport equation for all N. However, only the spreading front solutions are valid due to the formation of the shock for compressing fronts. Note that the same solution appears in Wu et al (1997: Equation 11b) expressed in travel distance, z.

A special simplification occurs for $N = 2$: $C(z,t) = C_0\left(\frac{Vt/z - 1}{\lambda N}\right)^{-1}$, which is a linear ramp from $C = 1$ to $C = 0$.

In the instance of a compressing front, applying the above solutions to the wave equation would lead to a non-physical solution that contravenes mass balance considerations and also predicts $\frac{C}{C_0} > 1$ (Figure 5.22).

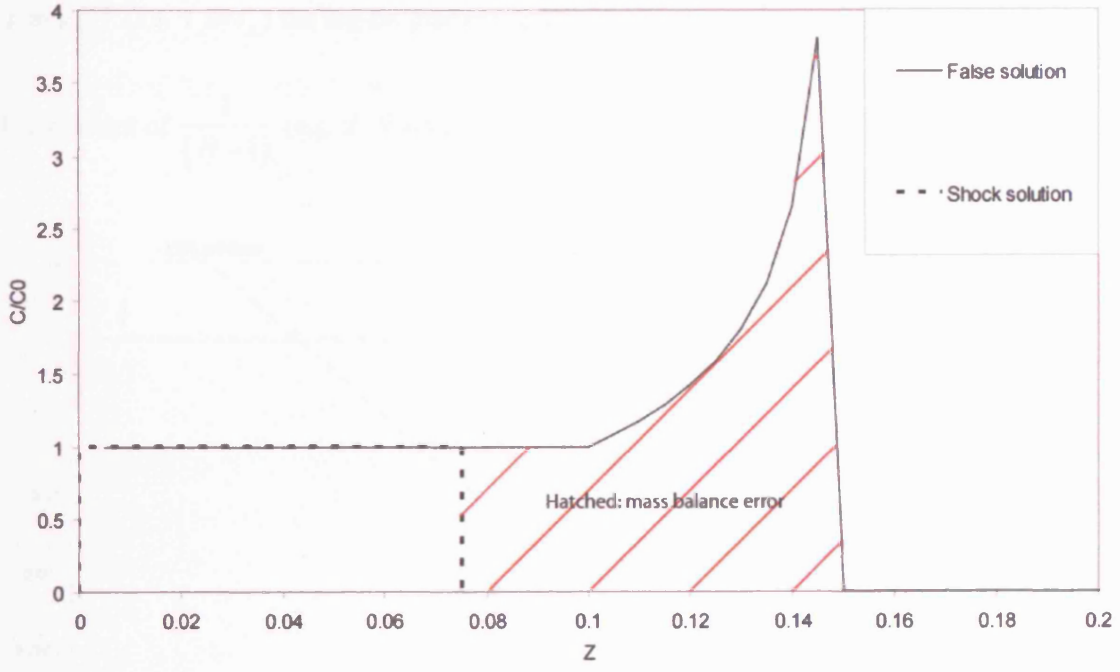


Figure 5.22: Shock solution compared to incorrect application of the analytical solution for step input, with Freundlich sorption, with $V = 0.3$, $\frac{B_d K}{\theta} = 0.5$, $N = 0.5$

5.4.5 Late-time gradients

The solution to Equation (5.44) for flushing when $N < 1$ and with a step input boundary condition $C(z,0) = C_I$, $C(0,t) = 0$ is:

$$C(z,t)=\left(\frac{Vt/z-1}{\lambda N}\right)^{1/(N-1)} \quad \text{for } t<(1+\lambda N C_I^{N-1})z/V \text{, and}$$
$$C(z,t)=C_I \text{ for } t<(1+\lambda N C_I^{N-1})z/V \tag{5.47}$$

Note that only one condition is needed for t for Equation (5.47), whereas two conditions were needed for Equation (5.46). The late time log-lin concentration gradient of Equation (5.47) is given by:

$$\frac{\partial}{\partial t}\log\left[C_I\left(\frac{Vt/z-1}{\lambda N}\right)^{1/(N-1)}\right]=\frac{V}{(1-N)(z-Vt)} \tag{5.48}$$

For $t \gg z/V$ (i.e. $t \gg t_A$) the log-lin gradient approximates to $\frac{1}{(N-1)t}$, which is of course a log-log gradient of $\frac{1}{(N-1)}$ (e.g. if $N=0.5$, the log-log gradient tends to -2). See Figure 5.23.

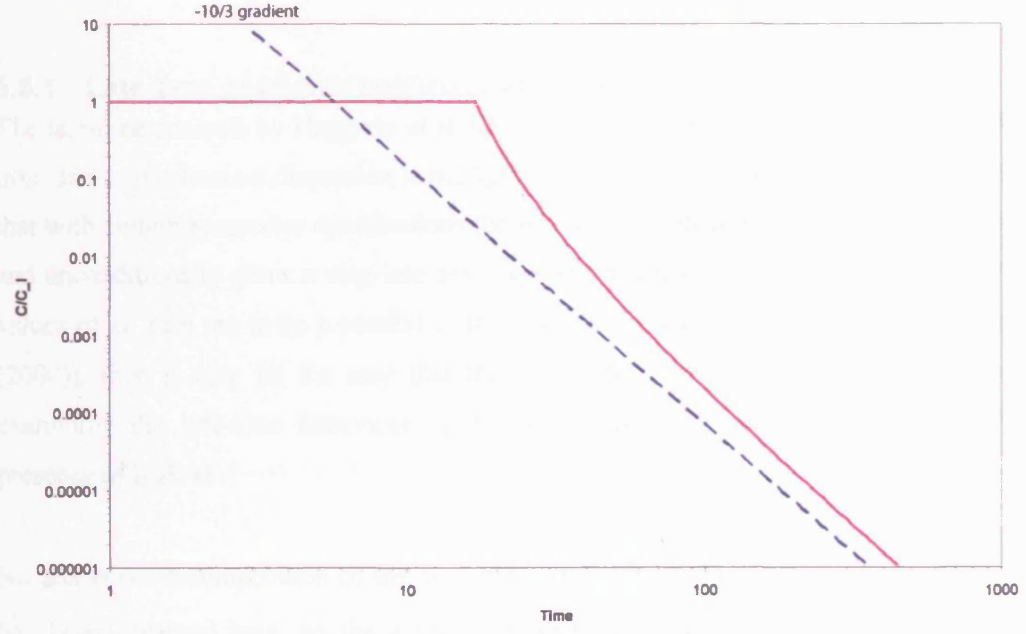


Figure 5.23: Late-time flushing with Freundlich isotherm and no mechanical dispersion ($N=0.7$, $\alpha=1$, $V=10$, $z=100$ and log-log gradient tends to -10/3)

It is worth noting that Appelo and Postma (1994:450) demonstrated the effect by comparing the flushing of a zero mechanical dispersion system of a linearly sorbed solute and a non-linearly sorbed one. The concentration is expressed in terms of pore volumes.

5.4.6 Moments

Bosma et al. (1996) stated in relation to BTCs for systems with non-linear sorption that, "no analytically expressions are yet available for the development of the various moments as a function of time". Numerical measures of moments are always possible, but this loses the advantage of showing the contribution to different moments by different parameters. This has not been done here.

5.5 *AD-DP model*

In the same way that the AD, DP and sorption models have been individually analysed for their characteristic behaviours, here the coupled AD-DP model is examined. The basis for examination here is the impulse response of the AD-DP transport equation,

$$\overline{C}_{\delta}^{f} = (m_0 / V_m)\exp\left[\frac{V_m z}{2D}\left(1-\sqrt{1+\frac{4g(s)\alpha}{V_m}}\right)\right] \tag{5.49}$$

5.5.1 Late-time gradients applied to AD-DP model

The late-time analysis by Haggerty et al. (2000) presumed that the effect on concentration at late time due to mechanical dispersion is negligible. However, Section 5.2 of this thesis established that with certain parameter combinations the AD can theoretically give a log-log gradient of -3/2 and unconditionally gives a very-late time exponential decay (log-lin linear). If extremely large values of α turn out to be a possibility in some systems (as conjectured by Becker and Shapiro (2000)), then it may be the case that this assumption does not hold. It is therefore worth examining the late-time behaviour of the two processes in conjunction, particularly in the presence of high α.

No analytical decomposition of the asymptotic behaviour of $C(t)$ for the combined processes has been obtained here, so the results are based entirely on numerical output. It might be conjectured that such an analytical solution might descend into a sum of exponentials, in which case one will become dominant.

Examination of the log-lin behaviour for the AD process alone required C to be diluted to very low levels before matching of the gradient occurred. This meant that in many cases numerical

error set in before any linearity was observed (neither proving nor disproving its existence). This is a drawback in terms of understanding the functional relationship, but not in practical terms, because the point for these functions at which double precision LT solutions breakdown is of the order 10^{-15}, which is considerably below most practical dilution problems.

Figure 5.24 demonstrates that the effect of increasing α is to reduce the late-time log-lin gradient away from that predicted by the DP-only gradient and apparently towards the AD-only gradient. Notice how numerical error sets in before a stable gradient is established (Figure 5.25).

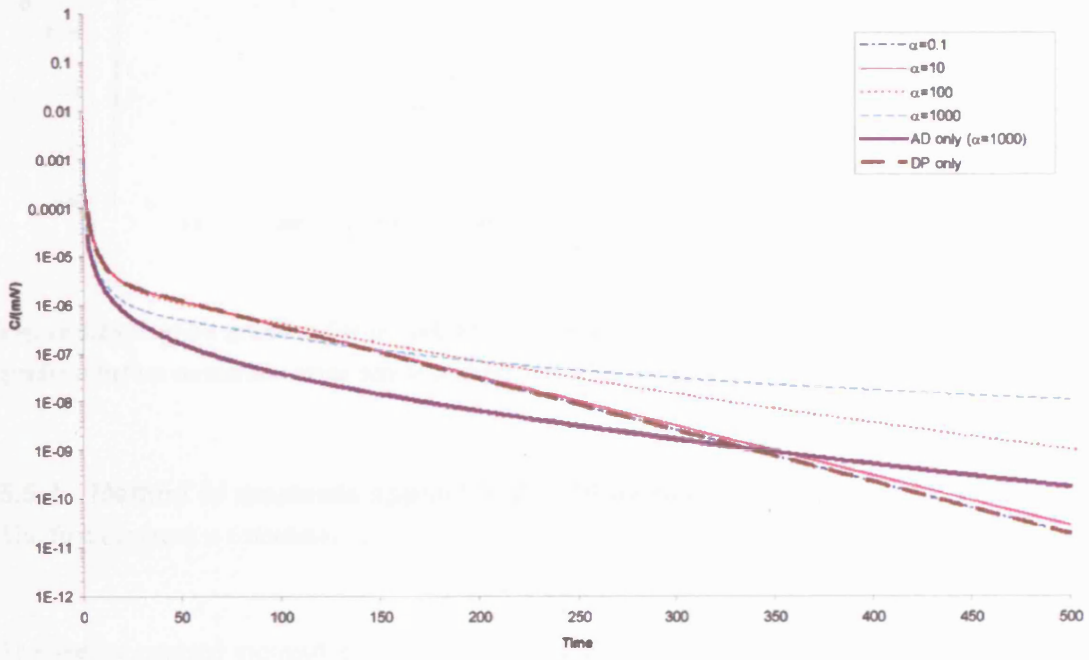


Figure 5.24: Log-lin plot of AD-DP model for $z = 0.05$, $t_{cb} = 100$, $\sigma = 5$, $\theta = 0.2$, $q = 1$, α =0.1, 10, 100 and 1,000 and with the extreme cases of AD only (for α =1000) and DP only.

With α =0.1 the log-lin gradient stabilises at -2.47×10^{-2}, which is equal to the late-time gradient predicted for a DP-only system, i.e. $-\pi^2/(4t_{cb})$. For α =10 a stable period of -2.37×10^{-2} appears. For α =100, there isn't a stable period up until the code crashes, but at $t = 500$ the gradient is 1.23×10^{-2} (Figure 5.24). This is apparently approaching the gradient required for an AD-only system, i.e. $-V/(4\alpha)$. Small α only influences the early time and as α is increased gradually, later times are affected until for large α the gradient of the tail is eventually effected.

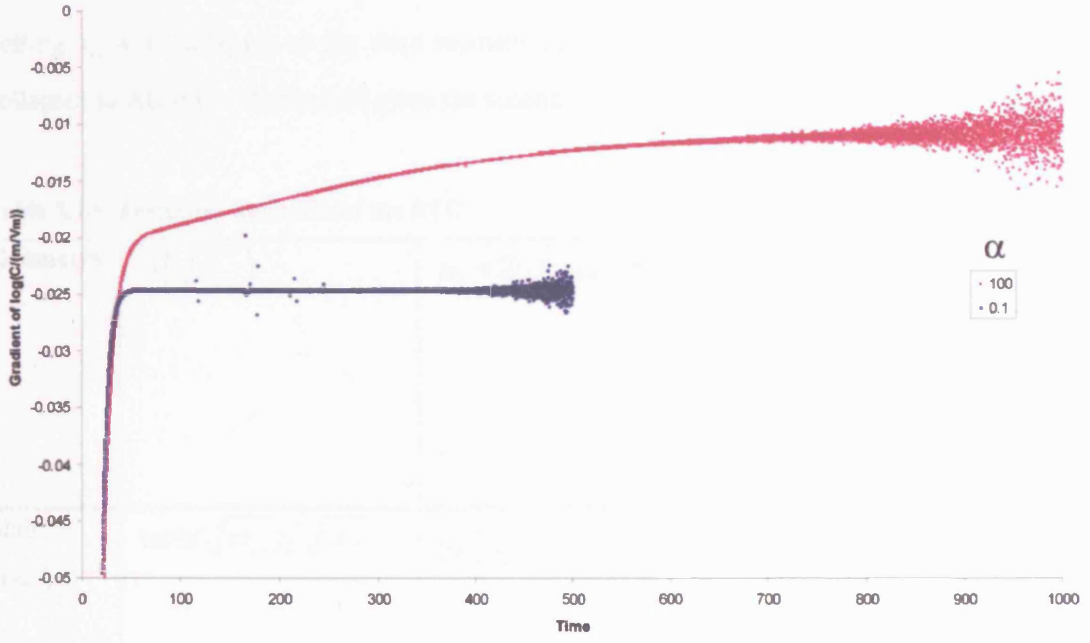


Figure 5.25: Log-lin gradient for α =100 and α =0.1 (for α =100, it does not settle to a constant gradient before numerical error sets in and for α =0.1 it settles to $\pi^2/(4t_{cb})$=2.47×10^{-2}).

5.5.2 Method of moments applied to AD-DP model

The first moment is calculated as

$$\mu_1' = t_a(1+\sigma) = t_A \tag{5.50}$$

The second centred moment can be decomposed into a linear contribution from the mechanical dispersive process and the DP process. So:

$$\mu_2 = \underbrace{2t_a t_D(1+\sigma)^2}_{AD} - \underbrace{2t_a\sigma\hat{B}'(0,t_{cb})}_{DP} \tag{5.51}$$

where $\hat{B}'(0,t_{cb}) = \frac{d}{ds}B\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$, $\hat{B}''(0,t_{cb}) = \frac{d^2}{ds^2}B\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$.

This could potentially be used as a method to decide which process dominates the dispersion of a solute.

The third moment shows a similarly useful decomposition, with the contributions being split into an AD term a DP term and a cross-term for the coupling of AD-DP. So:

$$\mu_3 = \underbrace{12t_a t_D^2(\sigma+1)^3}_{AD} - \underbrace{12t_a t_D \hat{B}'(0,t_{cb})\sigma(\sigma+1)}_{AD\text{-}DP} + \underbrace{3t_a \sigma \hat{B}''(0,t_{cb})}_{DP} \tag{5.52}$$

Setting $t_D = 0$ collapses to the third moment for DP-only and simultaneously setting $t_{cb} = 0$ collapses to AD only. Table 5.10 gives the second and third moments for common geometries.

Table 5.10: Temporal moments of the BTC

Geometry	$B\left(\sqrt{st_{cb}}\right)$	$\mu_2 = 2t_a\left(t_D(1+\sigma)^2 - \sigma\hat{B}'(0,t_{cb})\right)$	$\mu_3 = \underbrace{12t_a t_D^2(\sigma+1)^3}_{AD} - \underbrace{12t_a t_D \hat{B}'(0,t_{cb})\sigma(\sigma+1)}_{AD\text{-}DP} + \underbrace{3t_a \sigma \hat{B}''(0,t_{cb})}_{DP}$
Slab $b = h$	$\tanh(\sqrt{st_{cb}})/\sqrt{st_{cb}}$	$2t_a\left(t_D(1+\sigma)^2 + \frac{\sigma t_{cb}}{3}\right)$	$\mu_3 = 12t_a t_D^2(\sigma+1)^3 + 4t_a t_D t_{cb}\sigma(\sigma+1) + \frac{4}{5}t_a t_{cb}^2\sigma$
Cylinder $b = r/2$	$\frac{I_1\left(2\sqrt{st_{cb}}\right)}{\sqrt{st_{cb}}I_0\left(2\sqrt{st_{cb}}\right)}$	$2t_a\left(t_D(1+\sigma)^2 + \frac{\sigma t_{cb}}{8}\right)$	$\mu_3 = 12t_a t_D^2(\sigma+1)^3 + \frac{3}{2}t_a t_D t_{cb}\sigma(\sigma+1) + \frac{1}{8}t_a t_{cb}^2\sigma$
Sphere $b = r/3$	$\coth\left(3\sqrt{st_{cb}}\right)/\sqrt{st_{cb}} - 1/(3st_{cb})$	$2t_a\left(t_D(1+\sigma)^2 + \frac{\sigma t_{cb}}{15}\right)$	$\mu_3 = 12t_a t_D^2(\sigma+1)^3 + \frac{4}{5}t_a t_D t_{cb}\sigma(\sigma+1) + \frac{4}{105}t_a t_{cb}^2\sigma$
Infinite matrix	$1/\sqrt{st_{cb}}$	∞	∞
MIM	$\frac{\kappa}{\theta_{im}s+\kappa}$	$2t_a\left(t_D(1+\sigma)^2 + \frac{\sigma\theta_{im}}{\alpha}\right)$	$\mu_3 = 12t_a t_D^2(\sigma+1)^3 + 12t_a t_D t_M\sigma(\sigma+1) + 6t_a t_M^2\sigma$

*** note the zeroth moment for all the BTC systems (including the infinite matrix case) is m_0 / V_m**

5.5.3 Fitting applied to AD-DP model

5.5.3.1 General observations

Before the AD-DP model is examined by the method of moments, late time analysis and Monte-Carlo methods, the effect of systematically changing α with a fixed t_{cb} is demonstrated (Figure 5.26).

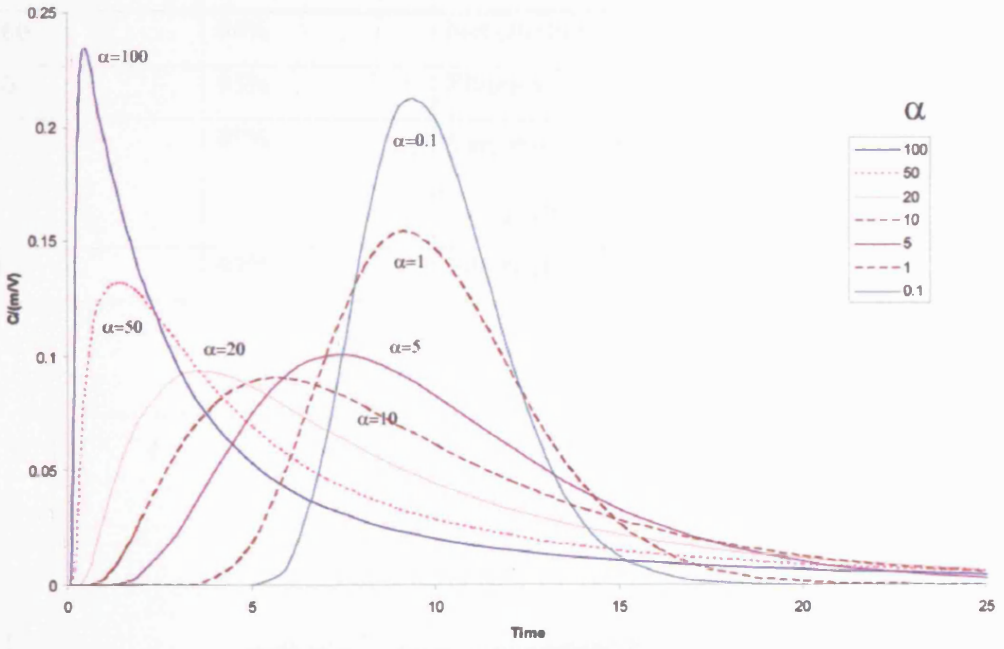


Figure 5.26: Unit impulse response for the AD-DP model, where $z = 50$, $t_{cb} = 1$, $\sigma = 1$, $\theta = 0.2$, $q = 1$ and α is varied.

5.5.3.2 Monte-Carlo applied to AD-DP model

The AD-DP model can be expressed in terms of three independent parameters (e.g. t_D, σ, t_{cb}). The results are shown as three 2D plots for each pair of parameters. The error-space for the AD-DP system is examined either side of the peak minimum that is shown in Figure 5.26.

Figure 5.26 showed that the minimum peak occurs at around $\alpha = 10$. A series of MC analyses were performed around this minimum, with synthetic noise added ($\sigma_N = 0.001(m_0 / V_m)$). The percentage of points within the predicted CR ellipsoid (for 5,000 realisations) are given in Table

5.11 along with a comment on the CR shape. The least elliptical representation was at α=1. The MC simulation for this worst case is shown in Figures 5.27-5.29. Notice how the ellipse has become noticeably curved and is narrow in the t_{cb} - t_D plane.

Table 5.11: Quality of the elliptical approximation for the AD-DP model for different α, where $z = 50$, $\theta = 0.2$, $q = 1$ and the targets are $t_{cb} = 1$, $\sigma = 1$

α	%	Comment
100	94%	Not elliptical, but close to 95%
10	95%	Elliptical
1	89%	Very thin in the t_{cb} - t_D plane and curved in σ - t_D plane
0.1	97%	Elliptical

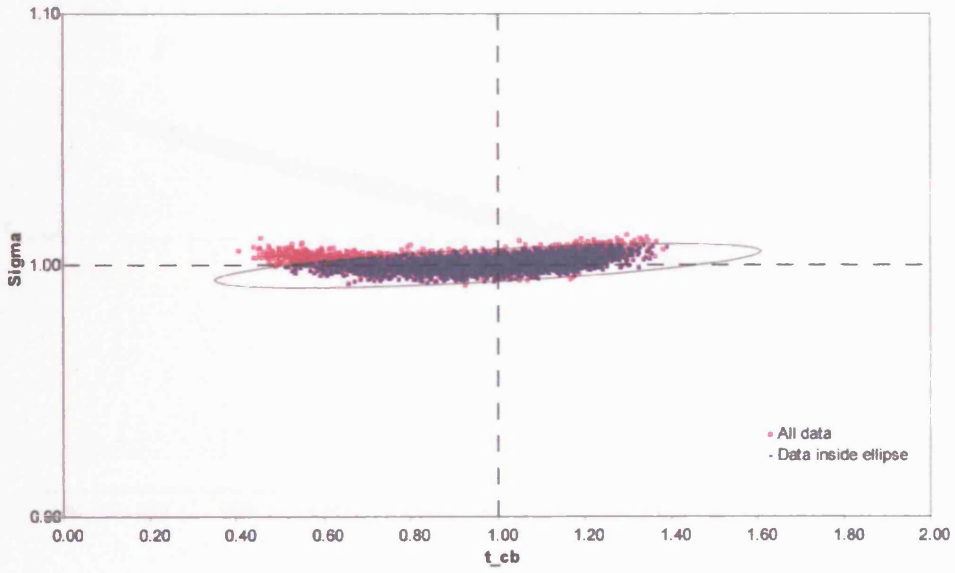


Figure 5.27: MC for AD-DP model in the t_{cb} - σ plane for target of σ =1, t_{cb} =1 and α =1 ($NMC = 10000$).

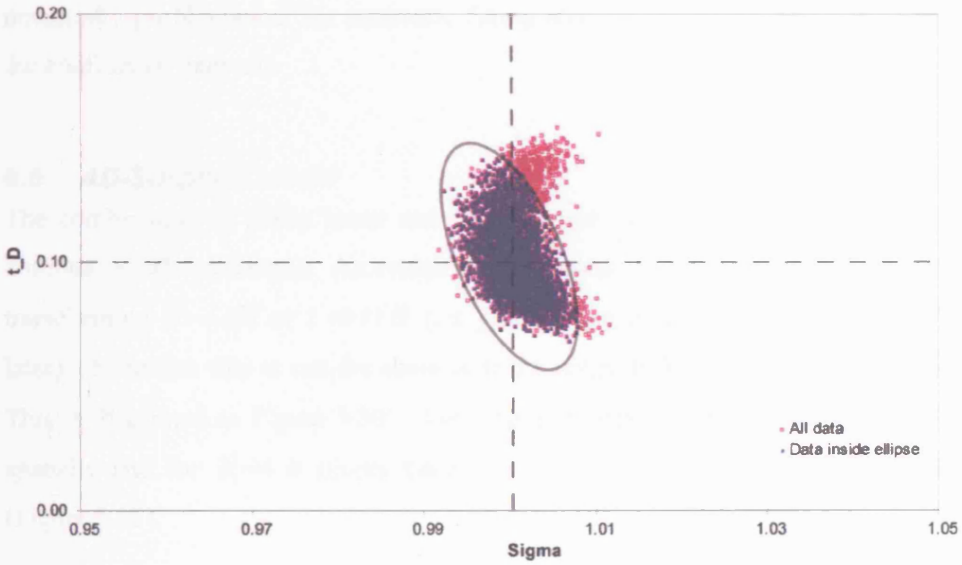


Figure 5.28: MC for AD-DP model in the σ - t_D plane for target of σ =1, t_{cb} =1 and α =1 ($NMC = 10000$)

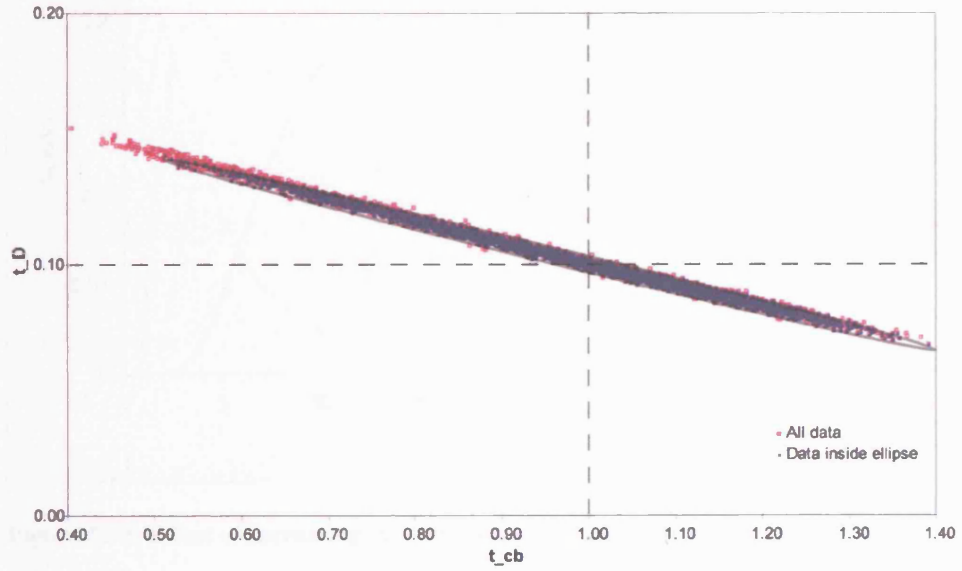


Figure 5.29: MC for AD-DP model in the t_{cb} - t_D plane for target of σ =1, t_{cb} =1 and α =1 ($NMC = 10000$)

The coupling between the three parameters of the AD-DP model creates for intermediate choices of α, curved t_{cb} - σ and σ - t_D planes and a strong line of non-uniqueness in the t_{cb} - t_D - plane. Taken together, the simple coupling of the two processes create an error space that is

potentially problematical for automatic fitting and for approximate linearised representations of the confidence intervals.

5.6 *AD-Sorption model*

The combination of firstly linear and secondly non-linear sorption and AD is examined (in the absence of DP exchange). As explained in Chapter 3, linear sorption has the simple effect of transforming $s \rightarrow sR$ or $t \rightarrow t/R$ (i.e. just shifting in time as if the whole BTC had occurred later). Note that this is not the same as if the entire BTC shifted to the right on the time axis. This is illustrated in Figure 5.30. Looking at it spatially, it is clear that R retards the plume spatially and for R=4 it simply takes four times as long to achieve the same spatial profile (Figure 5.31).

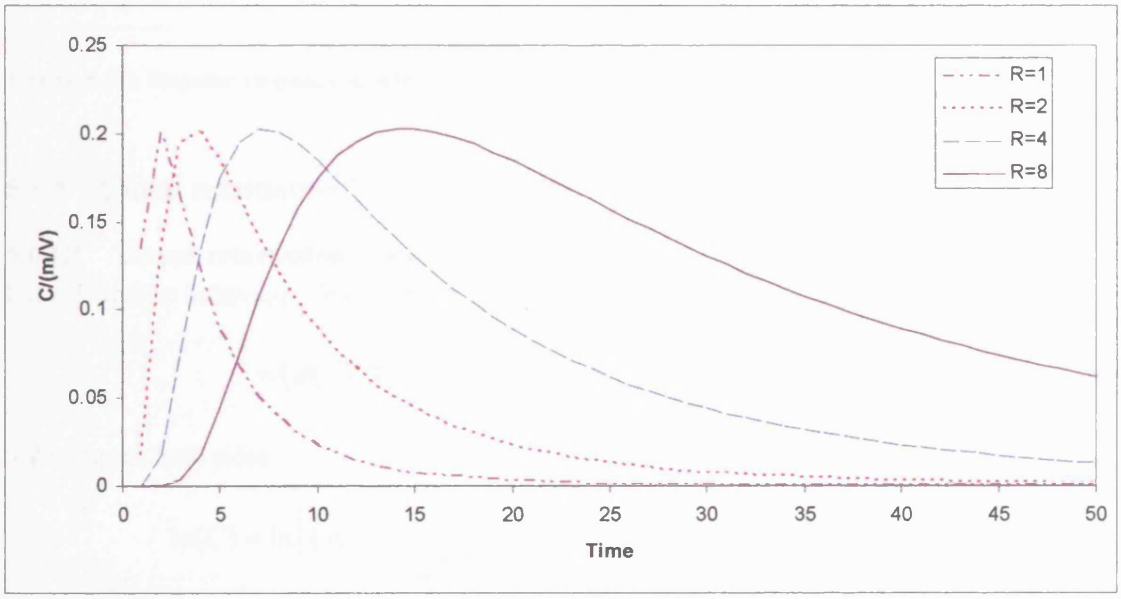


Figure 5.30: Effect of increasing R on impulse response to AD ($\alpha = 2, V = 1, z = 5$)

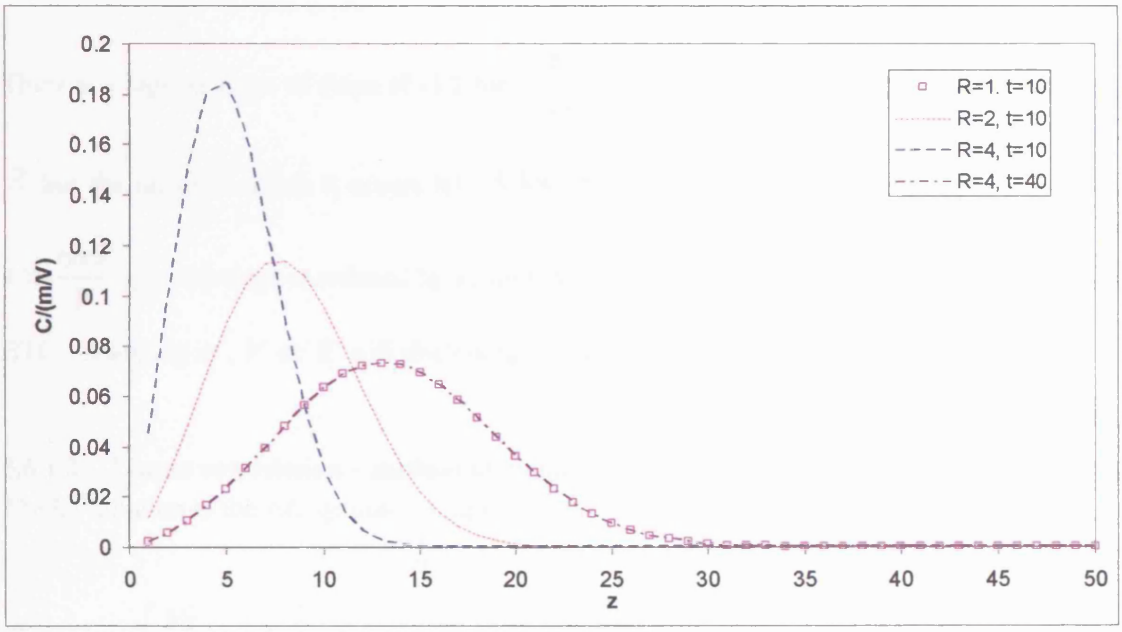


Figure 5.31: Impulse response to AD ($\alpha = 2, V = 1$), plotted in z

5.6.1 Linear retardation

5.6.1.1 Linear retardation – late-time

For asymptotic behaviour, first consider the semi-infinite AD impulse response:

$$C = (m_0 / V)\frac{z}{\sqrt{4\pi\alpha V t^3 / R^3}}\exp\left[-\frac{(z - Vt/R)^2}{4\alpha Vt/R}\right] \tag{5.53}$$

Take logs of both sides:

$$\ln(C) = \ln\left((m_0 / V)\frac{z}{\sqrt{4\pi\alpha V t^3 / R^3}}\exp\left[-\frac{(z - Vt/R)^2}{4\alpha Vt/R}\right]\right) \tag{5.54}$$

Differentiating gives:

$$\frac{\partial}{\partial t}\left[\ln(C)\right] = \frac{\partial}{\partial t}\left[\ln\left((m_0 / V)\frac{z}{\sqrt{4\pi\alpha V t^3 / R^3}}\exp\left[-\frac{(z - Vt/R)^2}{4\alpha Vt/R}\right]\right)\right] \tag{5.55}$$

$$\frac{\partial}{\partial t}\left[\ln(C)\right] = \left(\frac{Rz^2}{4\alpha V}\right)\frac{1}{t^2} - \left(\frac{3}{2}\right)\frac{1}{t} - \frac{V}{4\alpha R} \tag{5.56}$$

There is a log-log slope of slope of -3/2 for $\frac{Rz^2}{6\alpha V} \ll t \ll \frac{6\alpha R}{V}$ (i.e. the slope is unchanged by R but the range in which it occurs is). A log-lin gradient of $\frac{\partial}{\partial t}\left[\ln(C)\right] = -\frac{V}{4\alpha R}$ occurs for $t \gg \frac{6\alpha R}{V}$ (i.e. the slope is reduced by a factor R). Changing z will give an offset in the log-lin BTC. Changing α, V or R will also change the gradient.

5.6.1.2 Linear retardation – method of moments (MOM)

The LT solution to the AD system is simply:

$$\bar{C}_m = A\exp\left[\frac{Vz}{2D}(1-\zeta)\right] \text{ where } \zeta(s) = \sqrt{1+\frac{4g(s)D}{V^2}} \text{ and } g(s) = sR$$

By differentiation and taking limits, the moments of the ADE are produced and are given in Table 5.12.

Table 5.12: Moments of AD-sorption model.

Original	New (same R in both)
$\mu_1' = t_A$	$\mu_1' = Rt_A$
$\mu_2 = 2t_A t_D$	$\mu_2 = 2R^2 t_A t_D$
$\mu_3 = 12 t_A t_D^2$	$\mu_3 = 12R^3 t_A t_D^2$

It is striking that the Nth moments are now factored by R to the Nth power (i.e. the spread is increased as R^2, the skew is increased as R^3 and so on).

Comparing the first moments, an error in V could be made if a system that actually has a retardation, $R \neq 1$, is interpreted assuming $R = 1$:

$$\mu_1' = \frac{z}{V_1} = R\frac{z}{V_2} \tag{5.57}$$

Comparing the second moments, and simultaneously matching the first moments, gives no error in interpretation for dispersivity, α :

$$2t_{A1}t_{D1} = 2R^2 t_{A2}t_{D2} \text{ and } \frac{z\alpha_1}{V_1^2} = \frac{R^2\alpha_2}{R^2V_1^2}, \text{ hence } \alpha_1 = \alpha_2 \tag{5.58}$$

5.6.1.3 Linear retardation – fitting

An MC analysis of AD-linear sorption is not considered worthwhile since there is little substantive change from the AD-only model (except for the time shift). Well-behaved elliptical CRs are inevitable (as found in Section 5.2.4).

5.6.2 Non-Linear retardation

5.6.2.1 Nonlinear sorption (Freundlich)

Fetter (1999) presents a comparison of BTCs for $N = 1.3$ and $N = 0.7$ generated by using the Bio1D model[40] (Srinivasan and Mercer, 1987, 1988). This same example is used to test a new numerical code, SORBER.F90 (details in Appendix 1, page 437). This test, shown in Figure 5.32, demonstrates that $N = 1.3$ has a lower peak and arrives later than for $N = 0.7$.

[40] Note that the small typo in Fetter (1999:129) should read, "$B_d K_d$ value of 0.476 μg/g".

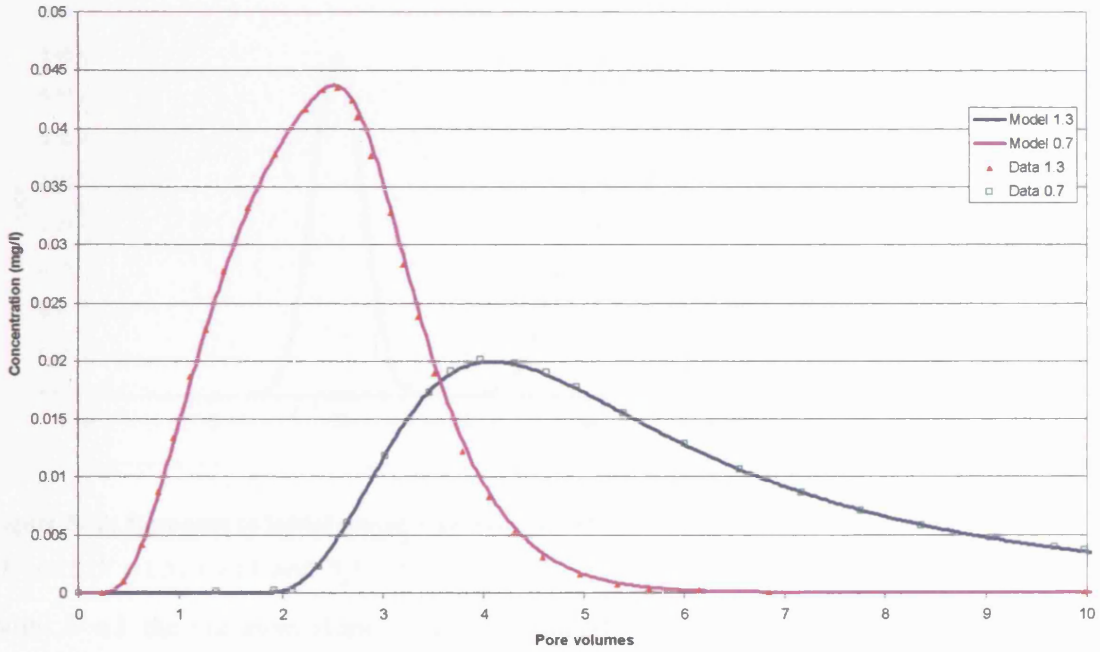


Figure 5.32: Test of SORBER.F90 against data from Grove and Stollenwerk (1984). $B_d K_d$ **=0.476 µg/g,** $D = 0.1$ **cm²/s,** $V = 0.1$ **cm/s,** $z = 8$ **cm,** $L = 16$ **cm,** θ **=0.37,** C_0 **=0.05 mg/l (note x-axis is pore volumes instead of time).**

It is interesting to see what is happening to the plume spatially, as well as considering BTCs. The underlying spatial symmetry is not always apparent when examining BTCs.

The effect of non-linear sorption on the spatial distribution of solute is given here, with the initial condition of a Gaussian spatial distribution of tracer.

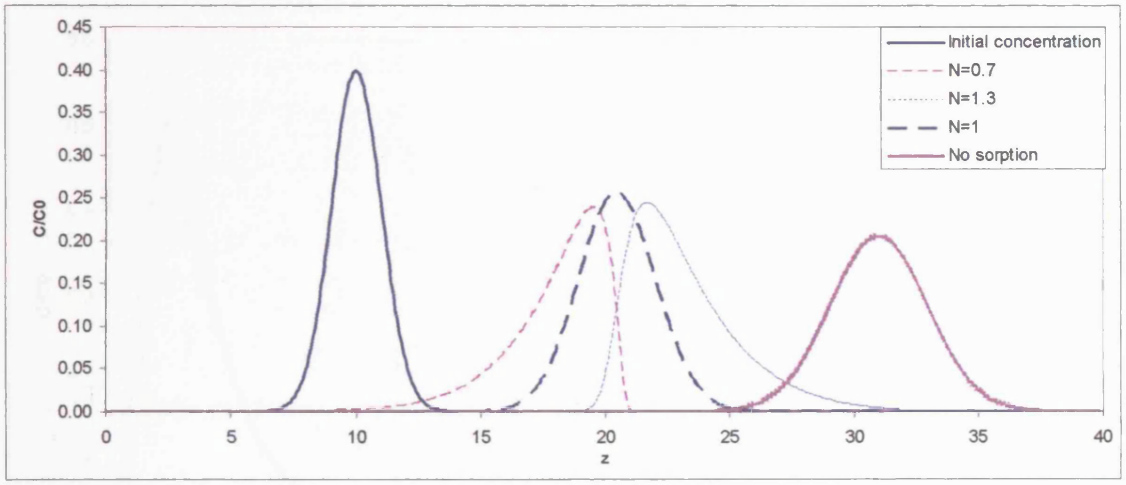


Figure 5.33: Response to initial spatial Gaussian distribution, $\mathrm{N}(10,1)$, with Freundlich isotherm ($D=0.1$, $V=1.5$, $t=14$ and $B_d K_d/\theta=1$)

With $N=1$ the Gaussian shape is retained spatially. For $N<1$ the plume is distorted by increasingly sharpening the leading edge and dispersing the trailing edge. The converse is true for $N>1$. The sharpened side continues to get steeper with time (and distance), until it becomes so steep that it is called a 'shock'. Note that with $N=1$ and $R=2$ the plume travels half the distance that it would have without sorption. The effect of nonlinear sorption on the peak is for $N<1$ to delay it relative to the $N=1$ peak and for $N>1$ to bring it forwards relative to the $N=1$ peak. But, by comparing to the plume with no sorption it is clear that it is not the case that nonlinear sorption can account for earlier peaks than would be expected for a conservative solute.

5.6.3 Non-linear - Late time

Because it is not clear from a theoretical point of view whether and under what conditions a Freundlich isotherm could be expected to remain representative, a distinct danger exists. This is of fitting very high concentration sorption data (for which the empirical fit to Freundich is due to saturation effects) and then using this at very low concentrations where the isotherm no longer applies. The only sure way to check this is to measure the isotherm down to the concentrations for which prediction is required. It is worth investigating how the Freundlich equation would influence the late time gradient if it turned out that it did hold down to low values of C. Literature searching revealed no relationship for asymptotic behaviour of C in the presence of Freundlich sorption and mechanical dispersion and one has not been derived here.

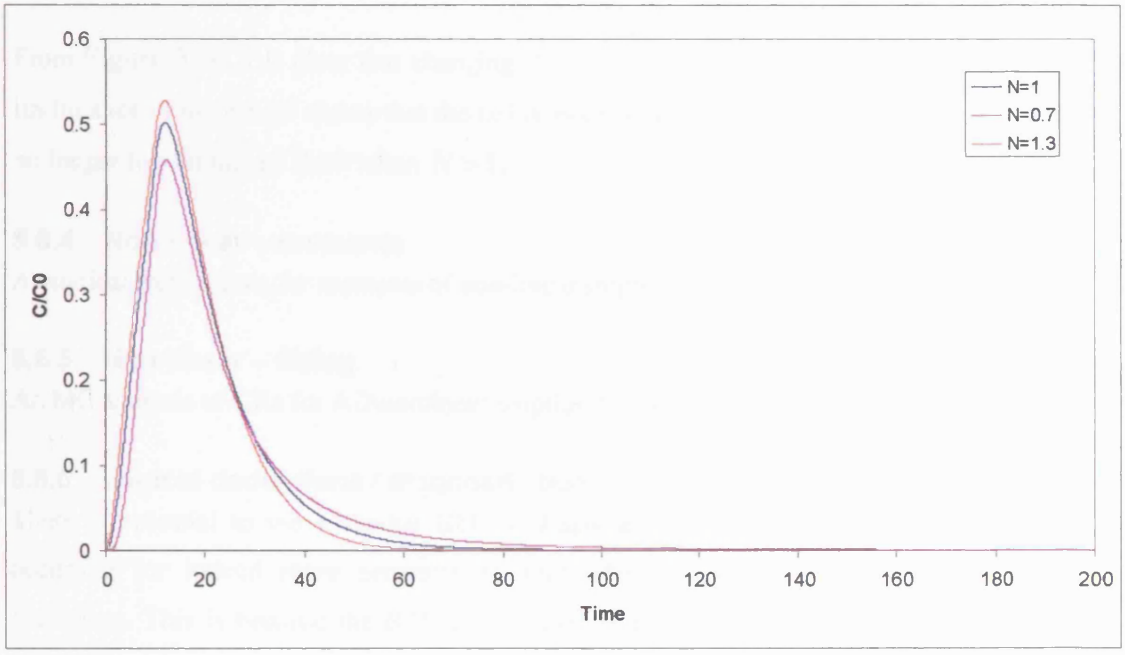


Figure 5.34: Arrival of solute compared for a top-hat pulse of injection $D = 10$, $V = 2$, $z = 10$, $T_p = 10$ and $B_d K_d / \theta = 1$ with Freundlich sorption.

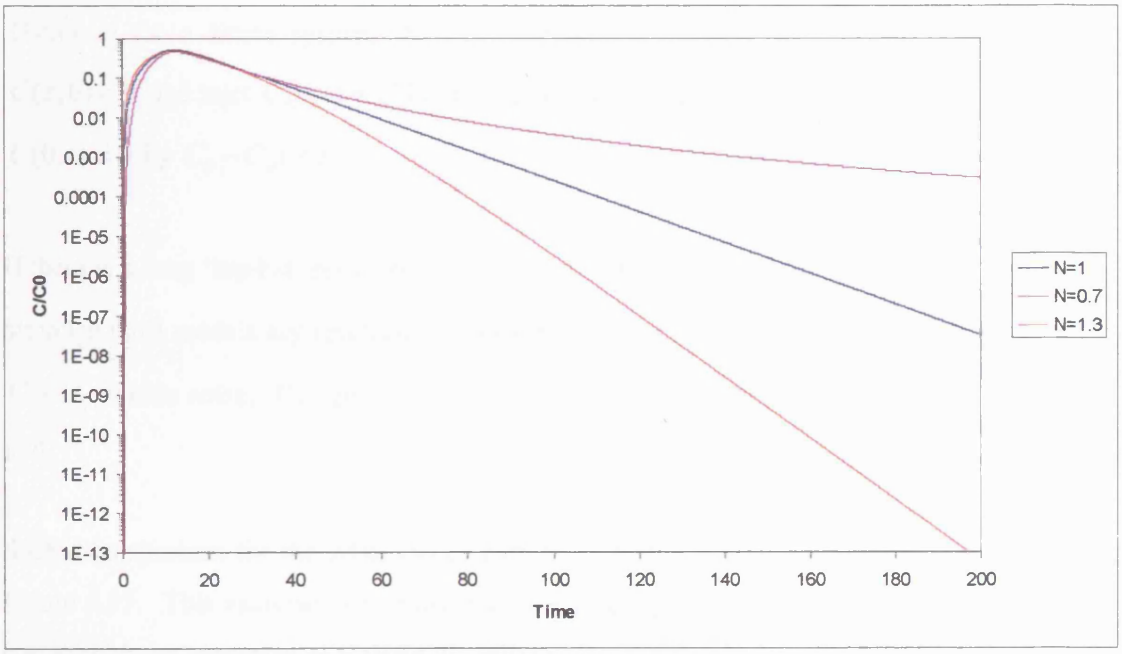


Figure 5.35: Arrival of solute compared for a top-hat pulse of injection $D = 10$, $V = 2$, $z = 10$, $T_p = 10$ and $\frac{B_d K_d}{\theta} = 1$ with Freundlich sorption (log-lin axes).

From Figure 5.34 it is clear that changing N has relatively little impact on a BTC viewed on lin-lin axes. Figure 5.37 shows that the tail is more substantially affected with a more rapid (and no longer log-lin linear) flush when $N > 1$.

5.6.4 Non-linear – moments

Analytical expressions for moments of non-linear sorption isotherms appear impossible.

5.6.5 Non-linear – fitting

An MC analysis of CRs for AD-nonlinear sorption has been left for future work.

5.6.6 Logical deductions / diagnostic test

There is potential to use a top-hat BTC to diagnose whether a non-linear sorption process is occurring (or indeed more generally to spot whether *any* non-linear transport process is occurring). This is because the BTC can be used to test for basic properties of linear systems: $\mathrm{F}(x_1(\tau)+x_2(\tau)) = \mathrm{F}(x_1(\tau)) + \mathrm{F}(x_2(\tau))$ and $\mathrm{F}(ax(\tau)) = a\mathrm{F}(x(\tau))$. Also, the condition for time-invariance (stationarity) is $y(t-\tau) = \mathrm{F}(x(t-\tau))$, where y is the output.

Therefore for a linear system, the step response, C_R, for a system with initial condition $C(z,0) = 0$ and inlet $C(0,t) = C_0\mathrm{H}(t)$ is related to the step response for $C(z,0) = C_0$ and inlet $C(0,t) = 0$ by $C_0 - C_R(x,t)$.

If there is a long 'top-hat' response in the presence of significant noise (say $\sigma_N = 0.01C_0$), then although most models asymptotically approach $C = C_0$ as $t \to \infty$, there comes a point at which $C \approx C_0$, within noise. This gives the possibility of comparing the passage from 0 to 1 and then 1 to 0.

Such a comparison for the ADE model with Freundlich sorption is plotted in Figure 5.36 and Figure 5.37. This exercise is repeated for $N = 1$ in Figure 5.38 and gives an identical pair of curves with the exception of a slight discrepancy at early time due to the fact that C is actually slightly less than 1 at the start.

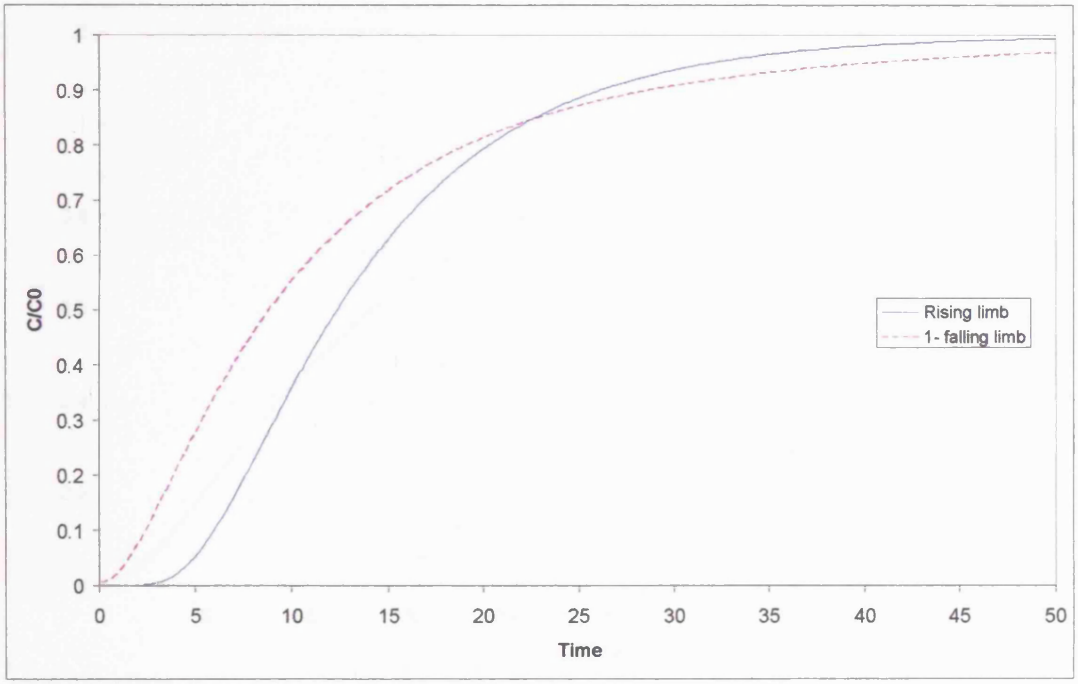


Figure 5.36: Rising and falling limbs compared for Freundlich sorption, $N = 0.5$, $D = 10$, $V = 2$, $z = 10$, $T_p = 50$ **and** $B_d K_d / \theta = 1$.

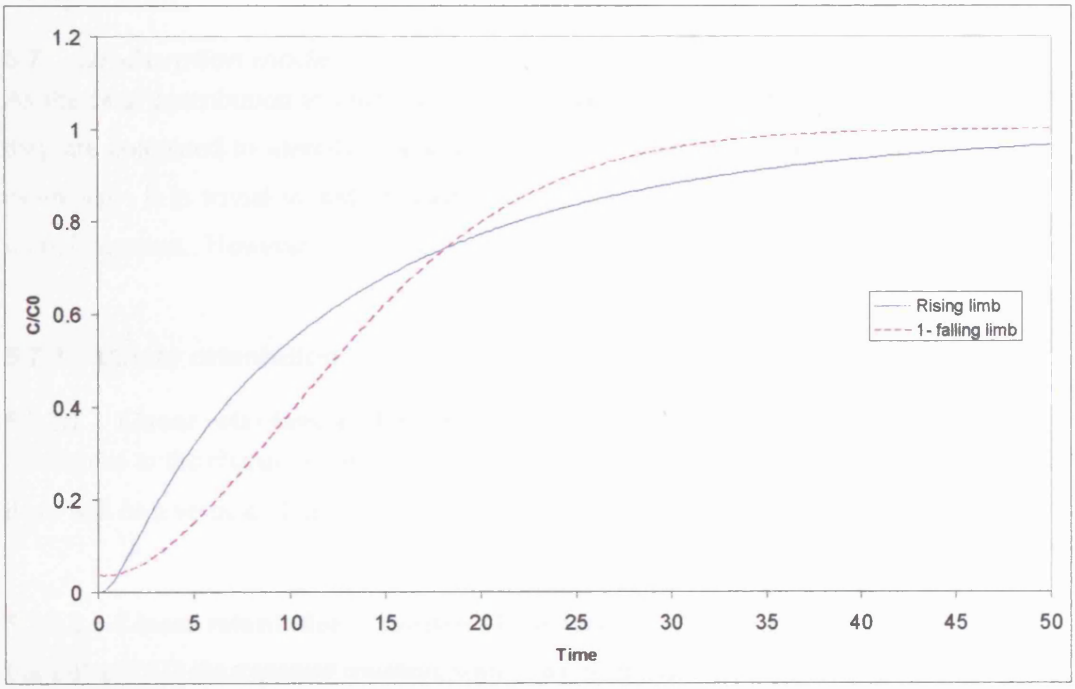


Figure 5.37: Rising and falling limbs compared for Freundlich sorption, $N = 2$, $D = 10$, $V = 2$, $z = 10$, $T_p = 50$ **and** $B_d K_d / \theta = 1$.

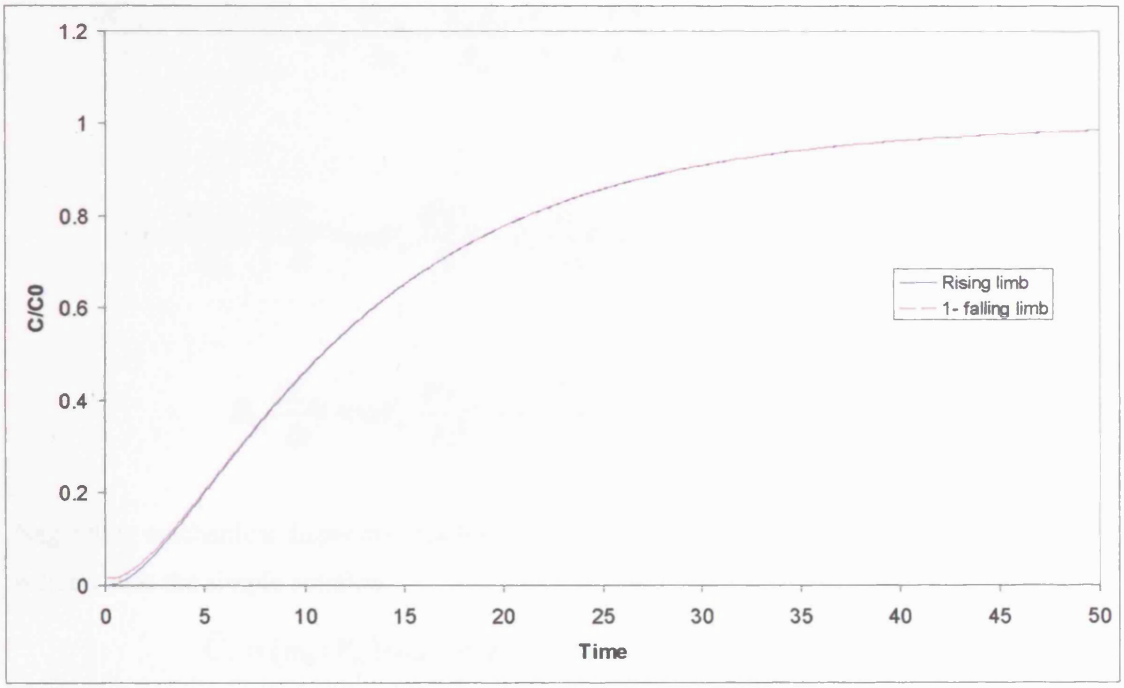


Figure 5.38: Rising and falling limbs compared for Freundlich sorption, $N = 1$, $D = 10$, $V = 2$, $z = 10$, $T_p = 50$ **and** $B_d K_d / \theta = 1$ **(note, lines overlap except at early time).**

5.7 *DP-Sorption model*

As the final contribution to objective 1, which needs a series of models to be appraised before they are compared to identify diagnostic traits, the coupling between DP and sorption is now examined. It is trivial to add in linear retardation, which may differ between the mobile and immobile zones. However, non-linear retardation is more difficult to model for the DP system.

5.7.1 Linear retardation

5.7.1.1 Linear retardation – Late time

Analogous to the change in the AD late-time system, the log-log gradient is unchanged (although there will be a vertical offset of the BTC) and the log-lin gradient is modified by a factor, R.

5.7.1.2 Linear retardation – method of moments

For DP and AD the transport equation, with linear sorption $C^* = K_d C$ is:

$$\frac{\partial C_m}{\partial t} + \underbrace{\sigma \frac{\partial \langle C_{im} \rangle}{\partial t}}_{q_e} + V_m \frac{\partial C_m}{\partial z} + \frac{K_d B_d}{\theta_m} \frac{\partial C_m}{\partial t} - \frac{K B_d}{\theta_{im}} \frac{\partial C_{im}}{\partial t} - \alpha V_m \frac{\partial^2 C_m}{\partial z^2} = 0 \tag{5.59}$$

$$\left(1 + \frac{K_d B_d}{\theta_m}\right) \frac{\partial C_m}{\partial t} = \alpha V_m \frac{\partial^2 C_m}{\partial z^2} - V_m \frac{\partial C_m}{\partial z} - \sigma \left(1 + \frac{K_d B_d}{\theta_{im}}\right) \frac{\partial \langle C_{im} \rangle}{\partial t} \tag{5.60}$$

$$R_m \frac{\partial C_m}{\partial t} = \alpha V_m \frac{\partial^2 C_m}{\partial z^2} - V_m \frac{\partial C_m}{\partial z} - \sigma R_{im} \frac{\partial \langle C_{im} \rangle}{\partial t} \tag{5.61}$$

Neglecting mechanical dispersion, taking LTs and solving in the instance of an impulse input of solute yields the simple solution:

$$\bar{C}_f = (m_0 / V_m) \exp\left[-t_a s \left(R_m + \sigma R_{im} \mathrm{B}\left(\sqrt{s R_{im} t_{cb}}\right)\right)\right] \tag{5.62}$$

Table 5.13 gives the moments for the DP-sorption.

Table 5.13: Comparison of moments for DP with DP plus linear sorption (see Table 5.7 for values of $\hat{\mathrm{B}}'(0)$ etc).

DP	DP-sorption (linear) (different R s)	(single R in both)
$\mu_1' = t_a(1+\sigma)$	$\mu_1' = t_a(R_m + R_{im}\sigma)$	$\mu_1' = t_a R(1+\sigma)$
$\mu_2 = -2t_a\sigma\hat{\mathrm{B}}'(0)$	$\mu_2 = -2t_a\sigma R_{im}\hat{\mathrm{B}}'(0,t_{cb})$	$\mu_2 = -2t_a\sigma R\hat{\mathrm{B}}'(0,t_{cb})$
$\mu_3 = 3t_a\sigma\hat{\mathrm{B}}''(0)$	$\mu_3 = 3t_a R_{im}\sigma\hat{\mathrm{B}}''(0,t_{cb})$	$\mu_3 = 3t_a R\sigma\hat{\mathrm{B}}''(0,t_{cb})$
$\mu_4 = 12t_a^2\sigma^2\hat{\mathrm{B}}'(0)^2 - 4t_a\sigma\hat{\mathrm{B}}'''(0)$	$\mu_4 = 12t_a^2 R_{im}^2\sigma^2\hat{\mathrm{B}}'(0,t_{cb})^2$ $-4t_a\sigma R_{im}\hat{\mathrm{B}}'''(0,t_{cb})$	$\mu_4 = 12t_a^2 R^2\sigma^2\hat{\mathrm{B}}'(0,t_{cb})^2$ $-4t_a\sigma R\hat{\mathrm{B}}'''(0,t_{cb})$

This is more general than the comparison give by Cunningham and Roberts (1998:1417), which is restricted in as much as the MIM transfer is modelled to be a kinetic sorption. Here, the MIM transfer is diffusive, but a retardation due to equilibrium sorption can be superposed.

Using $\theta R = \theta_m R_m + \theta_{im} R_{im}$ it can be shown that R_m and R_{im} do not have to be the same value in order to give the same first moment. Rearranging $\mu_1' = t_a \left(R_m + R_{im} \sigma \right)$ gives $\mu_1' = t_a \dfrac{\left(\theta_m R_m + \theta_{im} R_{im} \right)}{\theta_m} = t_a R(1+\sigma)$.

5.7.1.3 Linear retardation – fitting

An MC analysis of DP-linear sorption is not considered worthwhile since there is little substantive change from the DP-only model (except for the time shift).

5.7.2 Non-linear retardation

5.7.2.1 Non-linear retardation – Late time

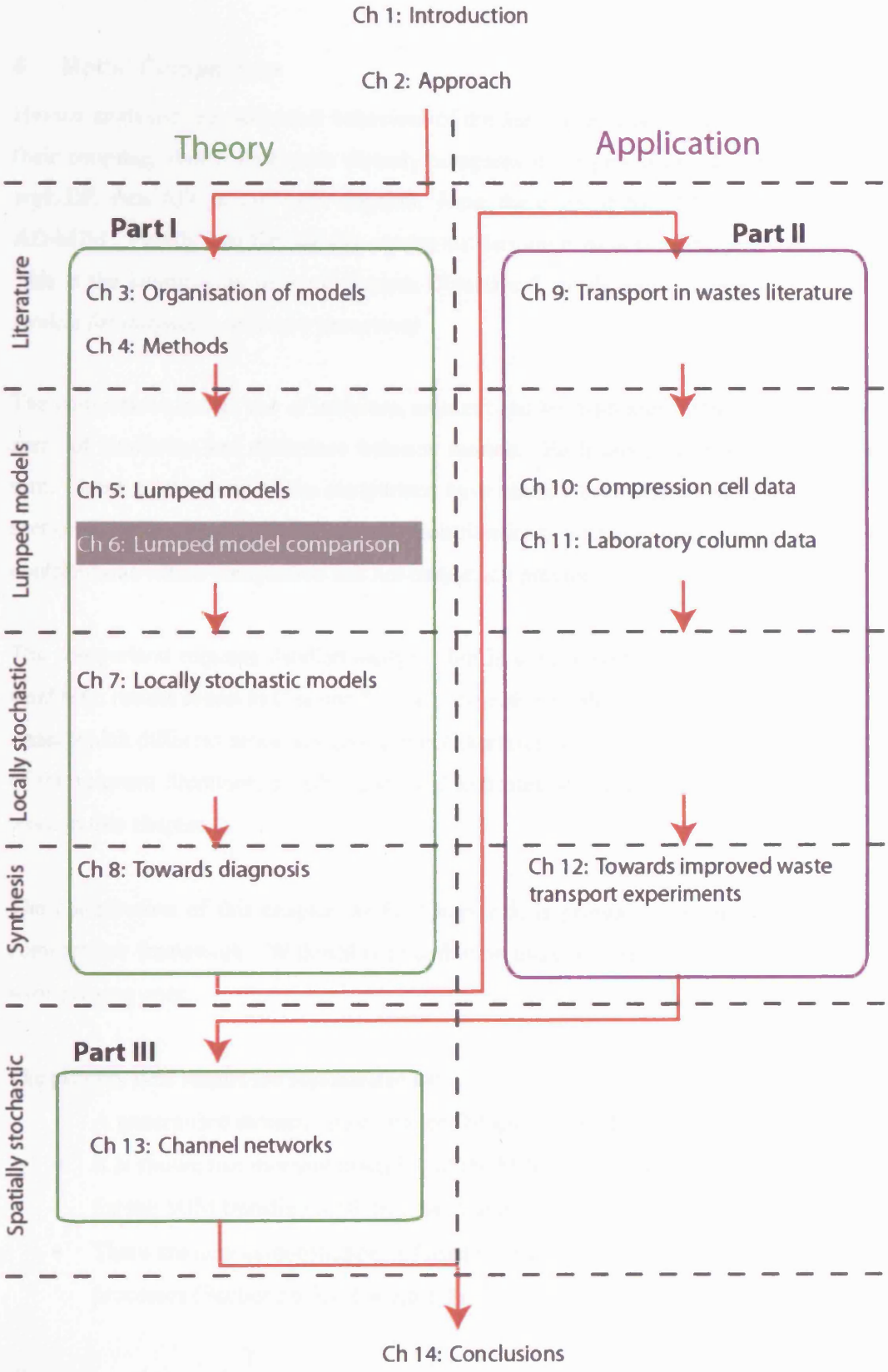
No code has been developed or obtained that can cope with non-linear sorption in the mobile and immobile phases in conjunction with AD in the mobile phase. This is therefore a task for the future.

5.7.2.2 Non-linear retardation – method of moments

It would be possible to take moments of the BTC for DP plus non-linear sorption numerically and might be worth exploring in the future.

5.7.2.3 Non-linear retardation – fitting

An MC analysis of AD-nonlinear sorption has been left for future work. Standard codes, such as Hydrus (Šimunek et al., 2005) or BIO1D (Srinivasan and Mercer, 1987) could perform this analysis for a MIM process, but no fully diffusive code with non-linear sorption is known.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

6 Model Comparison

Having analysed the individual behaviour of the lumped processes (AD, DP and sorption) and their coupling, this chapter now directly compares these processes. Firstly, AD is compared with DP, then AD (or DP) with sorption. Next, the coupled AD-DP process is compared with AD-MIM. Finally two DP transfer representations are compared (diffusive, i.e. DP) and MIM. This is the culmination of work towards Objective 2 *(to develop a comparative framework of models for diagnosis of simple processes).*

The comparison makes use of late-time, moment and least-squares methods, in order to map out areas of similarity and difference between models. Each comparison method is examined in turn. Some of the parts of the comparison have already been addressed in the literature. This section therefore organises the literature contributions, supplementing some and providing new contributions where comparison has not been made previously.

The comparison requires detailed analysis, but is summarised in Chapter 8 along with locally stochastic results found in Chapter 7. The hope is to be able to identify experimental conditions under which different processes give distinct characteristic behaviour. Table 6.1 identifies some of the relevant literature, possible gaps and indicates which are at least partially filled by the work in this chapter.

The contribution of this chapter, as for Chapter 5, is primarily the application of a rigorous comparative framework. Within this contribution there are inevitably new results interspersed with existing ones.

The primary new results are summarised as:

- A generalised moment matching for DP and AD models (Section 6.3.2)
- It is shown that moment matching of the MIM and DP BGFs results in a matched value for the MIM transfer coefficient that had previously been dismissed as 'inaccurate'.
- There are new demonstrations of least squares matching for a range of permutations of processes (Sections 6.3.3, 6.4.3,6.5.4)

The sections that are dominantly review material are 6.3.1, 6.3.3, 6.5,6.6.1 and 6.6.3.3.

Table 6.1: Summary of literature comparing simple process models for simple impulse or step input giving author and summary (*work needed is given in italics*, ✓= in this thesis, ×= future work)

Method	AD with DP	AD with AD-sorption	DP with Non-linear sorption	AD-DP with AD-MIM	DP with MIM (for BGF as well as BTC)
Late-time gradients	Becker and Shapiro (2000) suggested that data from Mirror Lake shows evidence of non-diffusive (i.e. AD) tailing. *Little done on this but conventionally DP is known to exhibit 'long' tails relative to AD.* ✓	*Late-time comparison of AD & AD-sorption is not in the literature.* ✓	*This work is needed.* ×	Haggerty et al. (2000) showed the late time gradients for the combined system. *Needs to be brought into common nomenclature.* ✓	Mathias. (2003) matched the first eigenvalues. He was primarily concerned with matching the BGF rather than the BTC. *The BTC as well BGF response needs matching.* ✓
Moments	Sanchez-Vila and Carrera (2004) compared moments for a limited range of geometries. *Geometry is generalised to include the MIM model and other geometries.* ✓	Valocci (1985, 1990) included linear retardation, R for moments of AD, DP and DP-MIM. *There are analytical difficulties for moments with non-linear sorption. This is not strongly needed , but could be done numerically. X*	*This is not greatly needed, but could be done numerically.* ×	Goltz and Roberts (1987) compared a limited set of DP geometries with MIM and AD-only. Valocci (1985) compared a spherical aggregate with MIM. Sardin et al. (1991) presented the moments for AD-MIM (without comparison). *Geometry is generalised. .* ✓	Barker (1985a) and Gerke and van Genutchen (1993) matched the BGFs by truncating series expansions. *The match points represented by these authors are shown to be moment-matches.* ✓ Cunningham and Roberts (1998) compared the DP(spherical) to MIM BGFs. *Geometry is generalised.* ✓
Least-squares (with noise model and weighting considered)	Bajracharya and Barry (1997) made the comparison using a spatially-stochastic AD model. *Simple AD and DP comparison by least-squares.* ✓ *Comparison when noise is added would be useful future work.* ×	Visual comparison of Non-linear with linear (Brusseau (1995)). *This work is needed (but saved for future work).* ×	*This work is needed.* ×	Young and Ball (1995) compared AD-DP (spherical geometry) model without noise with an AD-MIM model. Young and Ball (1997) extended to cylindrical geometry. Bajracharya and Barry (1997) compared a 2D spatially-stochastic numerical model to an AD-MIM model. *Systematic MC comparison needed.* ×	*This work is needed.* ×

6.3 DP & AD Matching

6.3.1 Late-time matching DP & AD

The BTC for the AD model has a tail, but it is generally thought to be much steeper than for systems with DP occurring. However Becker and Shapiro (2000) suggest that it is possible to explain late-time tailing by mechanical dispersion, rather than by any exchange mechanism. It is therefore worth making the comparison between the late-time gradients of the DP and AD processes to see what parameter values would be required to make a match (Table 6.2).

Table 6.2: Late-time gradients compared for δ(t) input and conditions under which they apply.

Model	Gradient of ln(C) vs. ln(t)	Gradient of ln(C) vs. t	Condition on ln(C) vs. t	Condition on ln(C) vs. t
AD	-3/2	$-\frac{V}{4\alpha}$	$\frac{6\alpha}{V} \gg t \gg \frac{z^2}{6\alpha V}$ or $6t_D \gg t \gg \frac{5}{3}t_a$	$t \gg \frac{6\alpha}{V}$
MIM	-	$-\kappa/\theta_{im}$	-	$t \gg t_a$
DP (spherical)	-3/2	$-\pi^2/9t_{cb}$	$t_{cb} \gg t \gg t_a$	$t \gg t_{cb}$
DP (slab)	-3/2	$-\pi^2/4t_{cb}$	$t_{cb} \gg t \gg t_a$	$t \gg t_{cb}$

Table 6.2 shows that for all except the MIM model, the log-log linear gradient is -3/2 but only for a certain window.

The 'tell-tale' -3/2 gradient indicator therefore requires a few caveats. For the DP model it occurs for apparently 'infinite' matrix diffusion (i.e. when $t \ll t_{cb}$). In the context of a mixture of block sizes, the timescale of the smallest fraction must also be much larger than the time of measurement. For systems with very high mechanical dispersivity ($Pe \ll 6$), the AD is also predicted to give a -3/2 gradient. For less dispersive situations (most field environments) there is little chance of confusion between the models, and a -3/2 gradient is reliably diagnostic of an effectively 'infinite' matrix diffusion behaviour. Any changes in dispersivity, velocity, length or diffusion coefficient will potentially change the period in which the linear gradient occurs, but not the gradient itself. A multiple tracer test in a DP (or indeed AD) system would therefore not be expected to show any difference in log-log gradient. Becker and Shapiro (2000) observed that the log-log concentrations of three different tracers are all approximately the same (with a gradient of -2) and therefore concluded that the system cannot be DP. However, a log-

log linear gradient that is different to -3/2 simply implies that the simple 1D DP model is inadequate. More complex diffusive systems might account for different gradients (this is discussed in Chapter 7). The effect of geometry of flow is also worth investigation, particularly for pumped-radial flows or flows in 'non-integer dimension' systems.

The existence of a log-lin linear late time behaviour does not distinguish any of the models as they all predict it. The log-lin gradient can be matched, for example for the AD, MIM and DP(spherical) models to all match requires:

$$\frac{V}{4\alpha} = \frac{\kappa}{\theta_{im}} = \frac{\pi^2}{9t_{cb}} \tag{6.1}$$

6.3.2 Moment matching DP & AD

A development of this comparison is already given in Sanchez-Vila and Carrera (2004), where the DP model is matched to the AD model. They concluded that matching the MIM model would be useful as it could include kinetic sorption, but left this open for future work. They did not generalise to *any* block shape, instead used a BGF that is restricted to a limited family (including slabs, cylinders and spheres). A more general representation is developed here. A criterion is also developed here for the matching condition that could be used to design column tests in which differentiation between these two process models is required.

The first condition that Sanchez-Vila and Carrera (2004) impose is that the Darcy velocity is the same in both systems (i.e. $V\theta = V_m\theta_m$), where V is the linear velocity in the AD system. This condition is fixed for a column experiment, where the inlet water flux should be a known constant value.

Table 6.3 compares the first four moments. These moments can then be matched based on this table.[41]

[41] Sanchez-Vila and Carrera (2004) refer to this process as the 'identification' of parameters and moments

Table 6.3: First four moments of the AD and DP models, where $\hat{B}^{(n)}(0,t_{cb})=\frac{d^n}{ds^n}B\left(\sqrt{st_{cb}}\right)\Big|_{s=0}$. **See Tables 5.4 and 5.5 for specific values of** $\hat{B}^{(n)}$.

Moment	AD	DP
μ_1'	t_A	$t_A=t_a(1+\sigma)$
μ_2	$2t_At_D$	$-2t_a\sigma\hat{B}'(0,t_{cb})=$ $-2t_a\sigma t_{cb}\hat{B}'(0)$
μ_3	$12t_At_D^2$	$3t_a\sigma\hat{B}''(0,t_{cb})=3t_at_{cb}^2\sigma\hat{B}''(0)$
μ_4	$12t_A^2t_D^2+120t_At_D^3$	$12t_a^2\sigma^2\hat{B}'(0,t_{cb})^2-4t_a\sigma\hat{B}'''(0,t_{cb})$ $=12t_a^2\sigma^2t_{cb}^2\hat{B}'(0)^2-4t_a\sigma t_{cb}^3\hat{B}'''(0)$

The **first moment match** is guaranteed by the matching of the Darcy velocities. The **second moment match** is expressed as $t_At_D=-t_a\sigma\hat{B}'(0,t_{cb})$, so:

$$t_D=\frac{\alpha}{V}=-\frac{\sigma}{1+\sigma}\hat{B}'(0,t_{cb})=-\frac{\sigma}{1+\sigma}t_{cb}\hat{B}'(0) \tag{6.2}$$

Next the **third moment match** is $4t_At_D^2=t_a\sigma\hat{B}''(0,t_{cb})$, so:

$$t_D=\sqrt{\frac{\sigma}{(1+\sigma)}\frac{\hat{B}''(0,t_{cb})}{4}} \tag{6.3}$$

This is a second condition on t_D. Equating this to the expression for t_D found for the second moment, gives:

$$\sigma=\frac{\hat{B}''(0,t_{cb})}{4\hat{B}'(0,t_{cb})^2-\hat{B}''(0,t_{cb})} \tag{6.4}$$

Finally, the **forth moment match** cannot be achieved exactly if the first three match conditions are imposed. However, the fourth moment can be expressed for the DP model, in terms of the matched AD parameters that were established for the first three moments, so:

$$\mu_4=12t_A^2t_D{}^2+t_At_D^3\left(64\frac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}\right) \tag{6.5}$$

These match points are the same as the results presented in Sanchez-Vila and Carrera (2004), but the BGF is now generalised.[42] Comparing this with the fourth moment for the AD model, $\mu_4=12t_A^2t_D{}^2+120t_At_D{}^3$, shows that the first ($t_A^2$) term is matched, but the second (t_A)

[42] Note that there is an error in Sanchez-Vila and Carrera (2004: Equation 19b) should read $\varphi_m=3\varphi/5$ instead of $\varphi_m=3\varphi_f/5$. This carries into Table 3 on page 171 of the paper, which is also incorrect.

term will have a variable quality of match depending on how close $\eta = 64\dfrac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}$ is to 120. Given that the value of η for the MIM model is the furthest from the value of 120 required for a perfect match, it is the *worst* match to the AD out of any of the BGFs considered here. A wider range of BGFs could now (in theory) be tested amongst which it is possible there is a geometry that gives a match for all four moments. How realistic such a geometry might be would remain to be seen. To match as far as the linear term in the fourth moment places the following constraint on the porosity ratio:

$$\sigma = \frac{\hat{B}''(0,t_{cb})}{4\hat{B}'(0,t_{cb})^2 - \hat{B}''(0,t_{cb})} \tag{6.6}$$

For certain systems, field knowledge may show that σ is considerably greater than is needed for a match. Combing the matches for the second and third moments gives:

$$t_D = \frac{\alpha}{V} = -\frac{\hat{B}''(0,t_{cb})}{4\hat{B}'(0,t_{cb})} = -t_{cb}\frac{\hat{B}''(0)}{4\hat{B}'(0)} \tag{6.7}$$

Therefore, $t_D \propto t_{cb} \propto \dfrac{1}{\kappa}$. It is interesting that the MIM exchange parameter, κ, scales with velocity, $\kappa \propto V/\alpha$. If the commonly observed scale effect for AD processes is included, $\alpha = 0.1z$, then $\kappa \propto V/z$. Likewise t_{cb} scales with velocity, as $t_{cb} \propto \alpha/V \propto z/V$

So, if an AD system is emulated by a MIM model (or vice versa), expect κ to increase with V and decrease with α (or with z if α/z is constant). If AD is emulated by DP, then t_{cb} will increase with z and decrease with V. A major point in recognising this is that it would be easy to make incorrect physical interpretations of such observed scaling. For example it might be concluded, following observation that t_{cb} apparently decreases with V, that increasing flow rate causes a more widespread flow organisation and effectively smaller immobile blocks but actually the system is described much more accurately by classical inter-grannular AD.

A different condition for good matching is when the first term in the fourth moment dominates over the 'incorrect' second term, so:

$$12t_A^2 t_D^{\,2} >> t_A t_D^{\,3}\left(64\frac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}\right) \tag{6.8}$$

which can be re-expressed as the condition, $Pe \gg \dfrac{16}{3}\dfrac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}$.

This is a slightly more meaningful condition than the condition given by Sanchez-Vila and Carrera (2004) that for a good match "this is valid only for large travel distance, where the linear term is less relevant". This is because there is no clear way of demonstrating what 'large' z really means ('large' after all is a relative term). Although, the two statements essentially come to the same thing for a known fixed α, since $Pe = z/\alpha$.

Table 6.4 evaluates this condition for a few geometries (giving approximately $Pe \gg 8$ for all the geometries). In the presence of the 'scale' effect which predicts $\alpha = 0.1z$ (therefore $Pe = 10$) it may be presumed that good matching of the fourth moment never occurs. However large Peclet numbers are conceivable, for example Sauty (1980) cited a range of $1 < Pe < 100$ as observed in real aquifers and Haggerty et al. (2000) cited a Pe of 1000. These larger Pe values potentially provide the condition for a good match for the first four moments.

In summary, the parameters of the match, σ, t_D and the conditions for a good match, η and Pe, are given in Table 6.4.

Table 6.4: Conditions that occur for moment matching of AD and DP.

	Parameters		**Conditions for good match**	
Model	σ	t_D	**Linear constant,** $\eta = 64\dfrac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}$	**Good fourth moment approximation condition,** $Pe \gg \dfrac{16}{3}\dfrac{\hat{B}'(0,t_{cb})\hat{B}'''(0,t_{cb})}{\hat{B}''(0,t_{cb})^2}$
MIM	1	$t_M/2$	$\eta = 96$	$Pe \gg 8.00$
Slab	3/2	$t_{cb}/5$	$\eta = 680/7 \approx 97.14$	$Pe \gg 170/21 \approx 8.10$
Sphere	5/2	$t_{cb}/21$	$\eta = 504/5 = 100.8$	$Pe \gg 42/5 = 8.4$
Cylinder	2	$t_{cb}/12$	$\eta = 99$	$Pe \gg 33/4 = 8.25$

Looking again at the second moment match, with the example of a slab and MIM models in comparison to the late-time match, a simple comparison is presented in Table 6.5.

Table 6.5: Comparison of AD with DP(slab) and AD with MIM

	AD & DP(slab)	AD & DP(sphere)	AD & MIM
Late-time semi-log match	$\frac{\alpha}{V} = \frac{1}{\pi^2} t_{cb}$	$\frac{\alpha}{V} = \frac{1}{4\pi^2} t_{cb}$	$\frac{\alpha}{V} = \frac{\theta_{im}}{4\kappa}$
Second moment match	$\frac{\alpha}{V} = \frac{\sigma}{1+\sigma} \frac{t_{cb}}{3}$	$\frac{\alpha}{V} = \frac{\sigma}{1+\sigma} \frac{t_{cb}}{8}$	$\frac{\alpha}{V} = \frac{\sigma}{1+\sigma} \frac{\theta_{im}}{\kappa}$

What is clear from Table 6.5 is that both forms of comparison give the same scaling between the parameters in that $\alpha/V \propto t_{cb}$ for DP and $\alpha/V \propto 1/\kappa$ for MIM.

6.3.3 Least-squares matching of DP and AD models

Bajracharya and Barry (1997) fitted an AD-MIM model to a 2D numerical model where grid-cells were randomly allocated a permeability (taken from a log-normal distribution). They showed that $\beta = \theta_m / \theta$ correlated with the variance in the log-permeability, σ_Y^2 in a negative exponential relationship. They did not explore the more simple least-squares comparison of AD and DP and did not introduce any noise in the simulated dataset. They found that there was either a linear or a power-law ($\kappa \propto V^b$) relationship. The linear relationship ($\kappa \propto V$) is adopted here since it was also found by Li et al. (1994) and Reedy et al. (1996). For a spatially stochastic permeability field, σ_Y^2 is proportional to the AD macrodispersivity. This is shown by Gelhar (1993) who relates the macrodispersivity, $\alpha_M = \sigma_Y^2 \lambda$, where λ is the integral scale (see Section 3.2.1.1.2). This can be combined to give a scaling of $\frac{\alpha_M}{V} \propto \frac{1}{\kappa}$ for fitting the AD-MIM to the AD model.

The three forms of matching are compared in Table 6.6, which extends Table 6.5 by including fitting.

Table 6.6: Comparison of AD with DP and AD with MIM

Match	AD & DP	AD & MIM
Late-time semi-log	$\frac{\alpha}{V} \propto t_{cb}$	$\frac{\alpha}{V} \propto \frac{1}{\kappa}$
Second moment	$\frac{\alpha}{V} \propto \beta t_{cb}$	$\frac{\alpha}{V} \propto \frac{\beta}{\kappa}$
Least-squares (Bajracharya and Barry, 1997; Li et al., 1994; Reedy et al., 1996.	-	$\frac{\alpha_M}{V} \propto \frac{1}{\kappa}$

A positive proportionality between β and variance in log-permeability is predicted by moments yet is not observed by fitting synthetic data or by late-time gradients. Table 6.6 also makes clear a proportionality between t_{cb} and α and an inverse proportionality between t_{cb} and V. It would be anticipated that the uncompleted entry in Table 6.6 would be $\frac{\alpha_M}{V} \propto t_{cb}$, but this would need to be confirmed by a fitting exercise in the same manner as those performed for matching AD and MIM.

Table 6.6 confirms that the same scaling relationships occur for matching of AD and DP (or MIM) processes for three different types of BTC comparison. It was not necessarily obvious without making the comparison that this would be true.

A single match for a DP(slab) synthetic dataset with an AD model is now demonstrated. A situation is chosen from the above moment match where the first three moments match.

The selected parameters were as follows; $z = 20$, $\sigma = 5/3$, $\theta = 0.1(1+\sigma)$, $\theta_m = 0.1$, $q = 1$, $V = q(\sigma+1)/\theta$, $\sigma_N = 0.005(m_0/V)$, $NMC = 10000$. Figure 6.1 shows the BTC as generated by the DP model, and the range of AD fits for all realisations within the 95% CR. The AD range brackets the DP BTC very closely. Figure 6.2 shows the MC scatter of best fits. This yielded a slightly curved distribution with a poor estimate for the 95% confidence region (77% of points are within the ellipse generated by the Hessian).

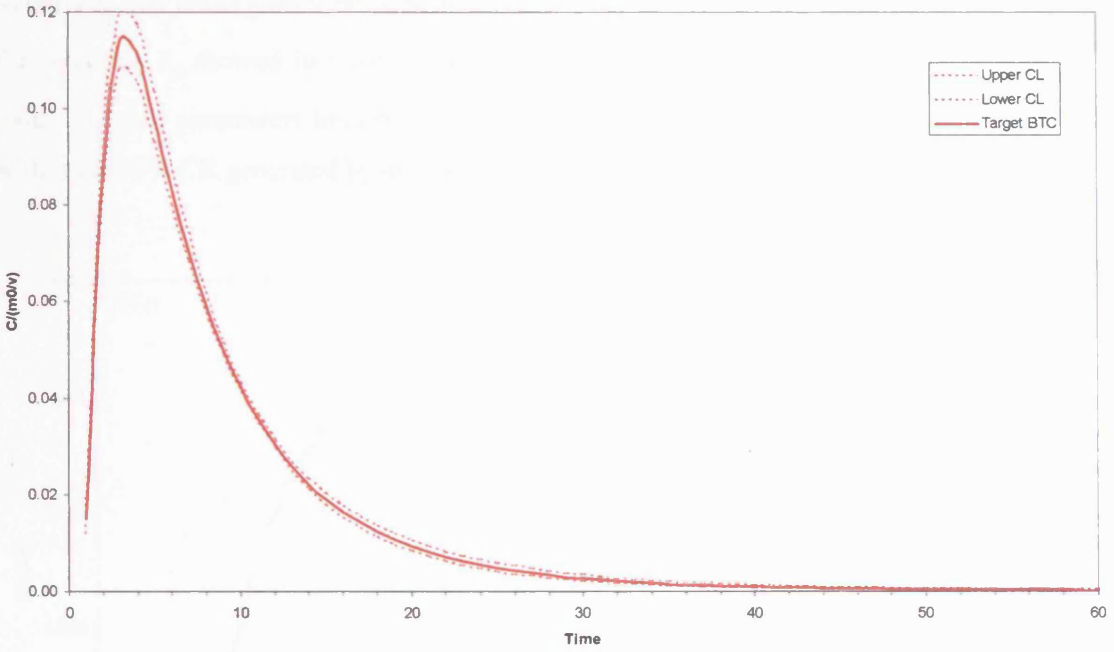


Figure 6.1: DP BTC (solid line) fit by the AD model (range at 95% confidence shown dotted) $t_{cb} = 10$

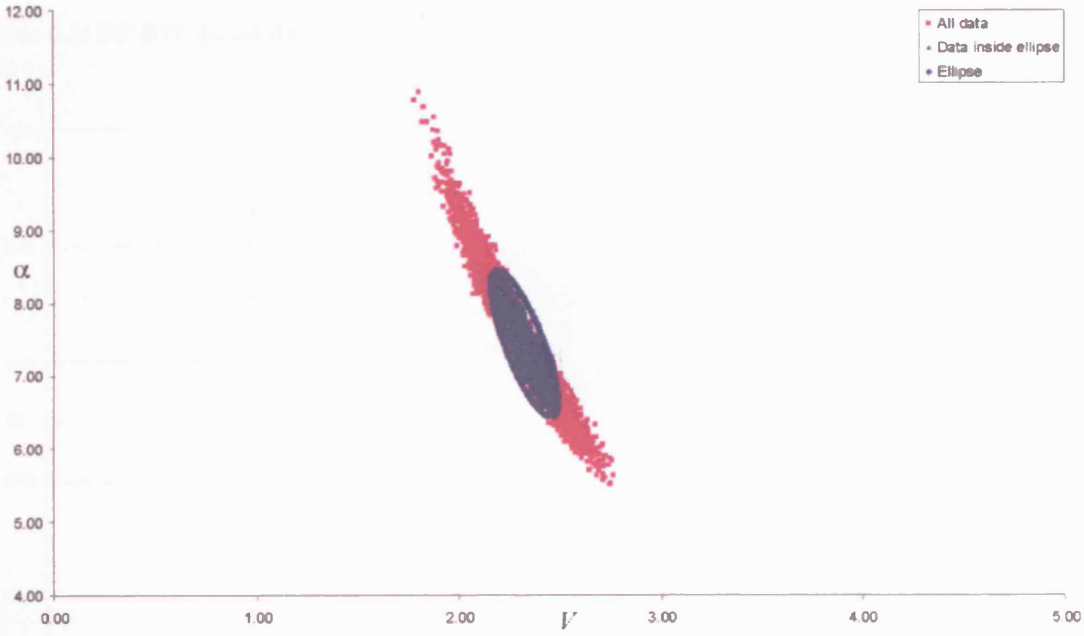


Figure 6.2: MC ellipse for V and α, $t_{cb} = 10$, where 7656/10000 points were inside the ellipse.

Relatively low noise gave a fairly wide CI in α, but less so for V. Examining the matching for reducing t_{cb} showed increasingly tight fits, for example Figure 6.3 demonstrates $t_{cb} = 1$ (with all other parameters held the same). The CR looks elliptical and 96% of the points lie within the 95% CR generated by the Hessian (Figure 6.4).

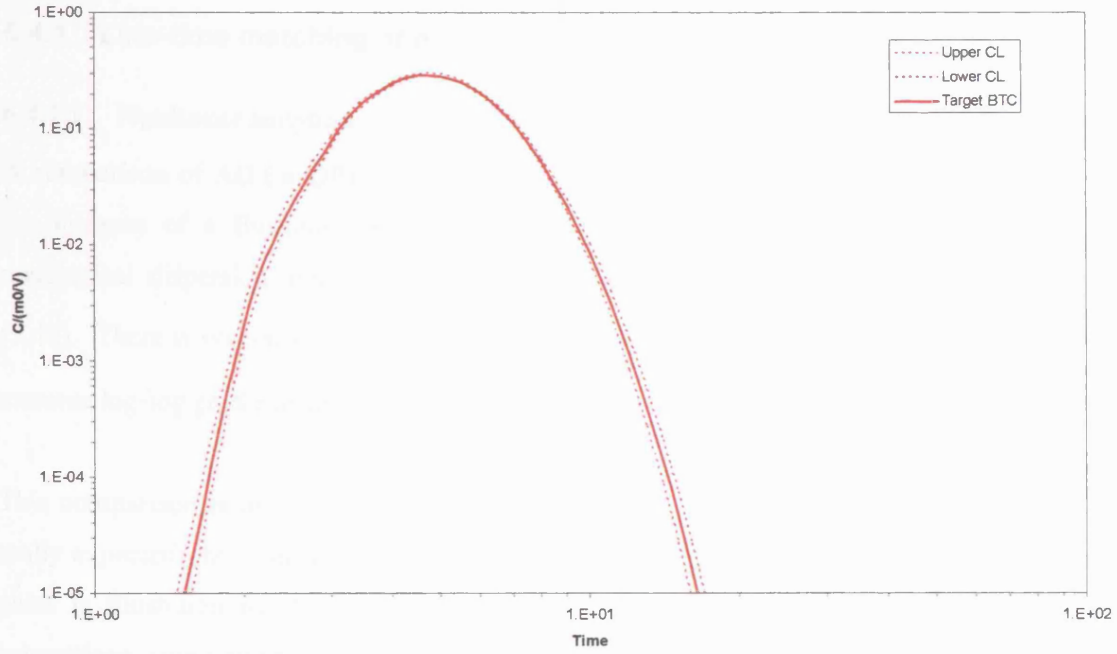


Figure 6.3: DP BTC (solid line) fit by the AD model (range at 95% confidence shown), log-log axes $t_{cb} = 1$

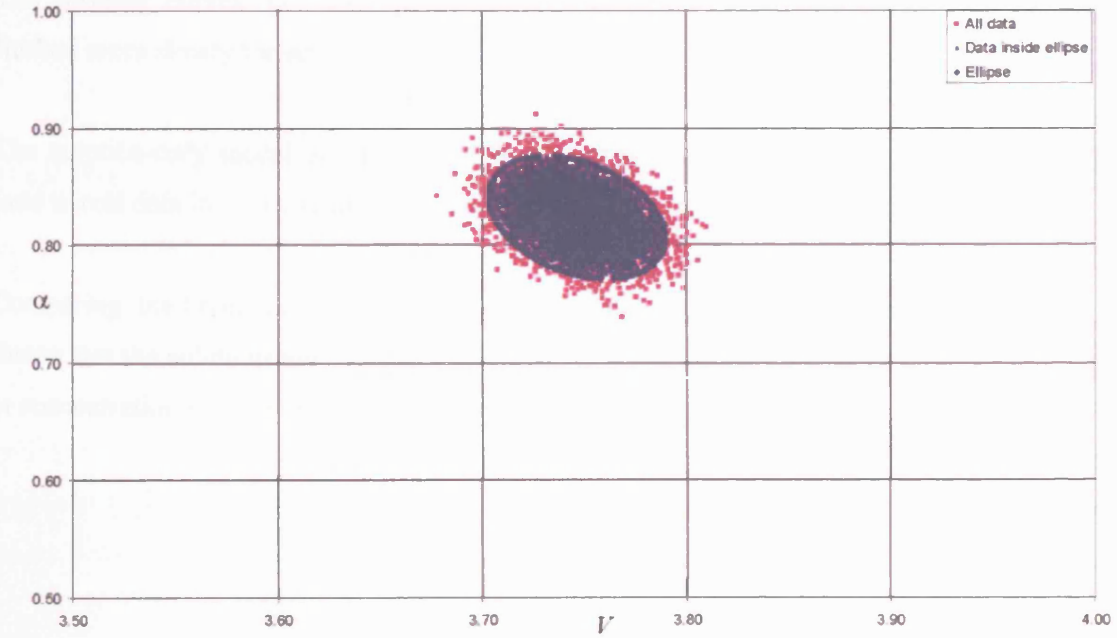


Figure 6.4: MC ellipse for V and α, where $t_{cb} = 1$, and 9556/10000 points were on or inside ellipse

This simple example shows that the CR can be both poorly and well described by calculation from the Hessian, depending on the underlying t_{cb}. There would need to be caution in plotting a wide distribution of best fits of AD fitted to DP as, in places, α can have quite a wide CI.

6.4 AD (or DP) and sorption models matching

6.4.1 Late-time matching of AD (or DP) and sorption models

6.4.1.1 Nonlinear sorption

A comparison of AD (or DP) and sorption in the absence of any other process can be made in the instance of a flushing system. The late time gradient of a flushing system with no mechanical dispersion and assuming Freundlich sorption (for $N<1$) is given by Equation (5.48). There is no constant log-lin gradient, but the log-log gradient tends to $1/(N-1)$. The extreme log-log gradients are -1 for $N=0$ and $-\infty$ for $N=1$.

This comparison is of course an 'end-member' and a complete absence of AD process is not really expected. In what sense the end-member is 'extreme' must be described carefully. This point is illustrated by the curves generated by Appelo and Postma (1994) who made the comparison, using number of pore-volumes as the x-axis. They plotted data, a numerical AD-sorption prediction and an AD-only prediction as well as the sorption-only analytical prediction. The sorption model predicted later arrival than AD-sorption, which in turn was later than AD. The flushing curves for AD-sorption and sorption cross over with the AD-sorption being flushed more slowly for larger pore volumes (or time).

The sorption-only model of Appelo and Postma (1994) was closer to the AD-sorption result (and to real data in a long sediment column) than the AD model.

Comparing the Freundlich flushing solution to an AD best-fit on a lin-lin time axis (Figure 6.5) shows that the solutions are not easily confused. The Freundlich BTC gives a sharp initial drop in concentration with a longer tail, whereas the AD BTC is more symmetric.

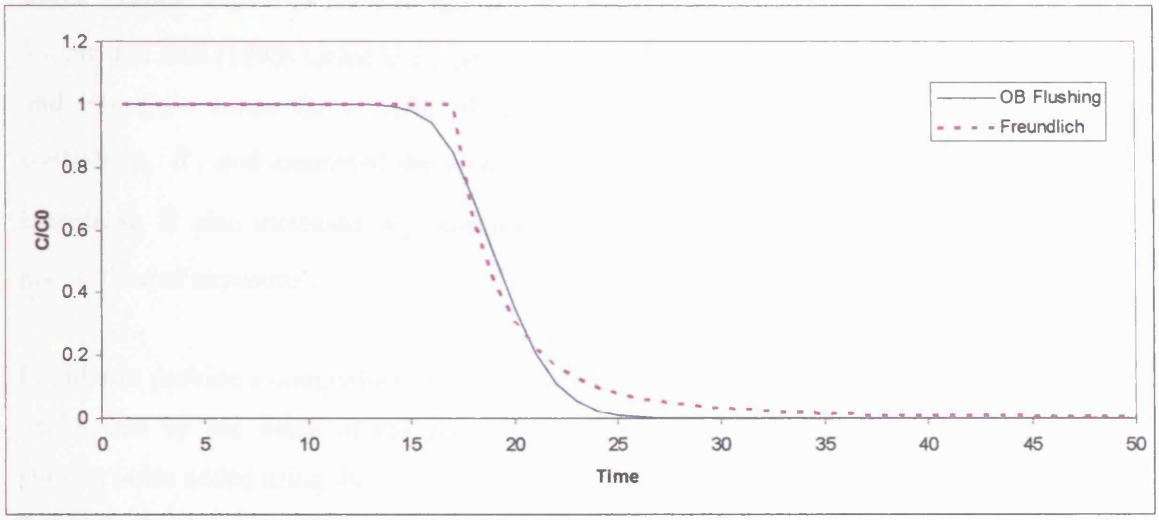


Figure 6.5: Flushing for the Ogata-Banks (OB) AD solution fitted to the pure Freundlich flush solution ($N = 0.7$, $\alpha = 1$, $V = 10$, $z = 100$ with AD fitting parameters $D = 3.49$, $V = 5.23$ and $z = 100$ fixed) .

The effect of non-linear sorption is illustrated in Figure 5.37 in Section 5.5.3. It makes the log-lin graph non-linear, curving away from the line described by pure AD. The findings for DP and sorption are similar.

6.4.1.2 Linear sorption

There is no effect of linear retardation, R, on the late-time log-log gradient for either DP-linear sorption or AD-linear sorption. Comparing the late-time log-log gradients therefore reveals nothing about R. R does, however, cause an offset in the late-time curve position in the y-direction. The log-lin gradients for DP-linear sorption and AD-linear sorption is the gradient of the unretarded process model, divided by R.

6.4.2 Moment matching AD (or DP) and sorption models

No moments have been obtained analytically for non-linear sorption.

Comparison can be made between AD and AD-linear sorption. Valocci (1985) included a retardation term when presenting moments. The key relationship from matching the first moment is that V becomes V/R. Applying this match to a comparison of the second moments gives no further relationships. The findings for DP are analogous but slightly revised by the immobile zone.

6.4.3 Least-squares matching AD (or DP) and sorption models

Young and Ball (1995) included a retardation term for linear sorption applied both in the mobile and immobile zones for a MIM model. Young and Ball (1995) varied the retardation coefficient, R, and examined the effect on dimensionless transfer coefficient, κ, viz. K_M. Increasing R also increased K_M because the system becomes further from equilibrium for a given 'time of exposure'.

In order to provide a comparison of a non-linear model to the basic AD and DP representations, uncluttered by the detail of the above study, some synthetic datasets have been generated without noise added using the simple numerical code developed here, SORBER.F90 (detailed in Appendix 1, page 438). Two inlet boundary conditions were tested - a step input of tracer and a 'top-hat' pulse. Least-squares fits of the AD, MIM and DP(sphere) models were then performed against this synthetic data (see Table 6.7 for the fit statistics and two fits are illustrated in Figures 6.6 and 6.7).

Table 6.7: r^2 over 2000 data points (spaced by time, Δt =0.1), with $T_p = 10$, $L = 10$, $D = 10$, $V = 2$, $N = 0.5$, $B_d K_D/\theta = 1$.

Input	AD	MIM	DP (sphere)
Step	0.9999	0.9311	0.9840
Pulse	0.9157	0.911	0.9840

It is possible to get a good fit if injection pulse duration, T_p, is allowed to float. In reality T_p at the inlet to a column, lysimeter or groundwater system is often well-known. But due to uncertainty over the inlet transfer function, perhaps due to mixing in the inlet or around a well, the effective T_p that the porous medium 'sees' is not the same. Holding T_p constant gave fairly poor fits to the top-hat input. Note the correlation coefficient yet again proved to be a potentially misleading measure, with the very high r^2 values quoted in Table 6.7 for MIM and DP corresponding to relatively poor visual fits.

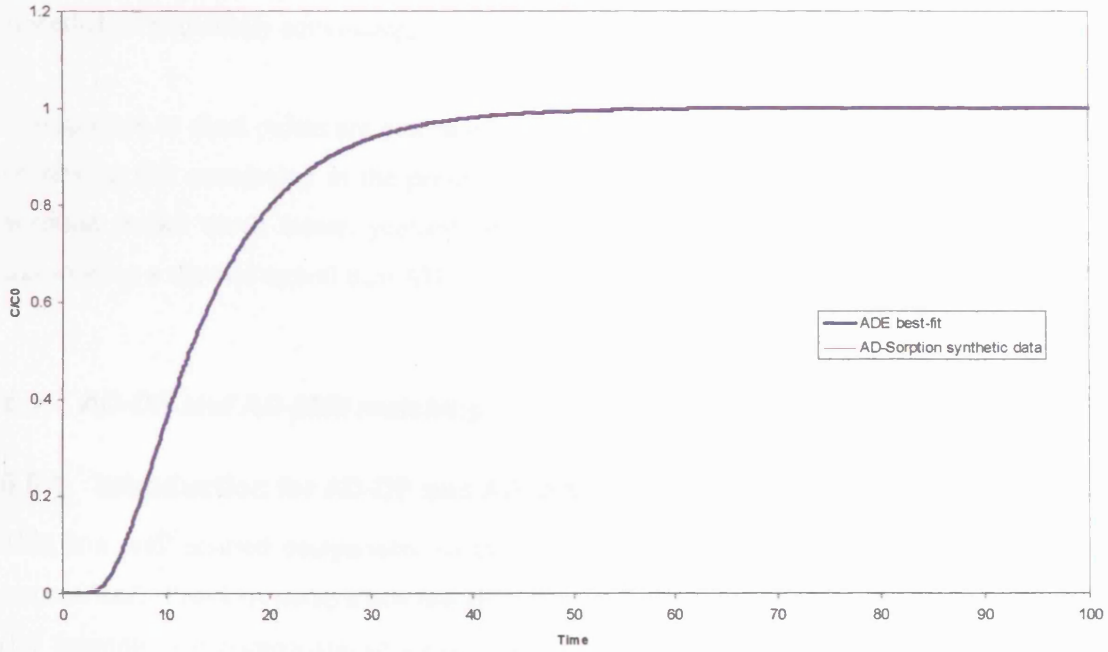


Figure 6.6: Near-perfect matching of AD-Sorption by an AD model, for step input with $T_p = 10$, $L = 10$, $D = 1.30$, $V = 0.68$ (R was allowed to float and the best fit was at $R = 1$).

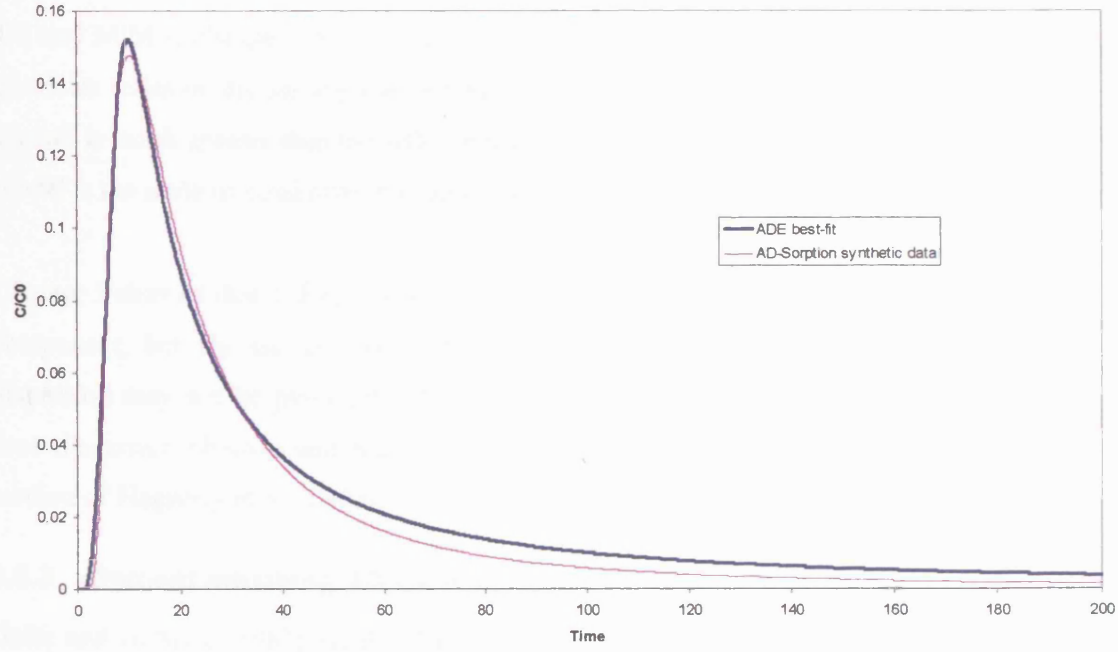


Figure 6.7: Best fit for AD-Sorption by an AD model, for top-hat input with $T_p = 0.41$, $L = 10$, $D = 2.03$, $V = 0.31$, $R = 1$

It may be impossible to tell the AD model from an AD-sorption model, in terms of their step responses, except possibly from examination of the parameter values or comparison to a known

conservative solute. The DP and MIM models compared to AD-sorption were poorer fits, but nonetheless potentially convincing, particularly in the presence of noise.

If responses to short pulses are compared, the difference in models is clear. It would be worth reviewing this conclusion in the presence of noise. The DP model was able to fit the AD-sorption model much better, probably because the DP process is more readily capable of introducing a skewed arrival than AD.

6.5 AD-DP and AD-MIM matching

6.5.1 Introduction for AD-DP and AD-MIM

This is a well studied comparison, so the results from earlier comparisons are gathered and summarised. Previous comparison has also been made more widely than this framework allows (for example, for comparison of sinusoidal inlet concentration by Zimmerman et al. (1993)). Chapter 8 further discusses intermittent flow.

6.5.2 Late-time matching AD-DP and AD-MIM

Haggerty et al. (2000) simulated the late-time behaviour of AD systems combined with different DP and MIM exchanges with a fixed $Pe = 1000$. Their analytical approximation for late-time gradients relies on the assumption that the DP component's contribution to the concentration in the tail is much greater than the AD component. In this case comparing the tail of the AD-DP model is the same as comparing the tail of the DP model.

Chapter 5 showed that if Pe is small enough then not only is the early time affected by the AD component, but the tail as well. However, it is possible that such levels of mechanical dispersion may not be physically plausible. The hypothesis of hydraulically-dominated late-time dispersion (Becker and Shapiro, 2000) is therefore untested by any application of the method of Haggerty et al. (2000).

6.5.3 Moment matching AD-DP & AD-MIM

Goltz and Roberts (1987) tabulated the moments for the AD-DP and AD-MIM models for comparison. Valocci (1985) compared the AD, AD-MIM and AD-DP(spherical), as well as AD-kinetic sorption. A linear retardation coefficient is retained in all the expressions. Cunningham and Roberts (1998) presented the same comparison. Sardin et al. (1991) also found moments for AD-MIM.

The results in Chapter 5 allow a more general comparison in terms of BGFs. These are recalled here, but with the different process components annotated:

$$\mu_1' = t_a(1+\sigma) \tag{6.9}$$

$$\mu_2 = \underbrace{2t_a t_d(1+\sigma)^2}_{AD} - \underbrace{2t_a\sigma\hat{B}'(0,t_{cb})}_{DP} \tag{6.10}$$

$$\mu_3 = \underbrace{12t_a t_d^2(\sigma+1)^3}_{AD} - \underbrace{12t_a t_d\hat{B}'(0,t_{cb})\sigma(\sigma+1)}_{AD-DP} + \underbrace{3t_a\sigma\hat{B}''(0,t_{cb})}_{DP} \tag{6.11}$$

These expressions for the first three moments can be used for specific comparisons given assumed geometries. For example between DP(sphere) and MIM (Table 6.8). Matching μ_2 and assuming that t_d is the same $\frac{t_{cb}}{15} = \frac{\theta_{im}}{\kappa}$ is the same as matching the first moment of the BGFs. μ_3 cannot be matched simultaneously with μ_2, assuming that t_d is the same.

Table 6.8: Second and third moments for AD-DP(sphere) and AD-MIM.

Geometry	μ_2	μ_3
Sphere	$2t_a t_d(1+\sigma)^2 + \frac{2}{15}t_a\sigma t_{cb}$	$12t_a t_d^2(\sigma+1)^3 + \frac{4}{5}t_a t_d t_{cb}\sigma(\sigma+1) + \frac{4}{105}t_a t_{cb}^2\sigma$
MIM	$2t_a t_d(1+\sigma)^2 + \frac{2}{\kappa}t_a\sigma\theta_{im}$	$12t_a t_d^2(\sigma+1)^3 + \frac{12}{\kappa}t_a t_d\theta_{im}\sigma(\sigma+1) + \frac{6}{\kappa^2}t_a\sigma\theta_{im}^2$

6.5.4 Least-squares matching AD-DP and AD-MIM

AD fit by AD-MIM

Bajracharya and Barry (1997) fitted an AD model by AD-MIM. They generated synthetic data based on the assumption of spatially stochastic K (and therefore macroscopic AD). A stochastically-centred view of reality is appealing and in line with the views of Gelhar (1993) who saw convective effects as dominating transport in most instances. The insinuation that could possibly be taken from such a comparison is that the AD-MIM model is therefore an empirical approximation to this 'reality'. However, reality *could* be more like an opposite case - an AD-DP system which may be approximated by the AD model.

AD-DP fit by AD-MIM

Young and Ball (1995) covered this comparison fairly comprehensively. Synthetic data were generated using an AD-DP(spherical) numerical model, without any addition of noise. The AD-MIM model was simulated and fitted using CXTFIT (van Genuchten, 1999), by varying at first only κ and then both κ and $\beta = \theta_m / \theta$ (although they remarked that the second form of fitting loses meaning). Pe was fixed at 400 in both models. They did not present the goodness-of-fit, simply stating the parameter values at which this best fit was found (making use of several non-dimensional groups). However, some of the fits that are presented are visibly poor. They plotted the dimensionless MIM coefficient, defined by $K_M = \kappa t_{cb} / \sigma t_a$, against the time of the experiment (the 'truncation time') for different $\gamma = \sigma t_a / t_{cb}$, showing: K_M reduces with longer truncation (this is because a high κ is needed to fit the early data); K_M reduces as γ increases (for the same reason); K_M reduces with a longer pulse time. It is the rather approximate measure of 'time of exposure' or 'contact time' of solid to the solute that they find to be the key issue in determining K_M. The length of step input they defined as the time of the experiment minus the mobile arrival time t_a normalised by t_{cb}. For a pulse it is the pulse time that was again normalised by t_{cb}.

They further plotted K_M against dimensionless truncation time and dimensionless pulse time. This gave a straight line with a negative gradient on a log-log axis (which curved off to a constant K_M at long time). A more reliable relationship between parameters of the different models was established when a step inlet boundary concentration was used. This is interesting, but unsurprising, since a 'top-hat' pulse would be expected to be a more demanding case for fitting (for example Section 6.4.3 showed how the step responses of different models can be readily fitted). The scatter may be explained by the shape of the error around the minimum. This lack of examination of the error space is a weakness and would make a useful further study.

Any change in the boundary conditions or system parameters that tend to increase the contact time, reduce K_M. Log-log plots of dimensionless κ against dimensionless time were linear for the earlier times and tend to the asymptotic exchange value for very long time. The primary conclusion of the work is that there is a danger that varying column conditions could be erroneously confused with differences in the intrinsic diffusion rate (or with t_{cb}).

Young and Ball (1997) further extended this analysis to cylindrical macropores surrounded by an annular region (which is the same as the 'hollow tube' model of Chapter 5). The conclusion

of both papers is that caution should be used in interpreting fitting results for κ, since changes in other variables (such as velocity, length, solute contact time) may be responsible for the changes rather than there being a true change in the rate of exchange.

In summary, Young and Ball (1995, 1997) covered the AD-DP(spherical) and AD-DP(hollow tube) – AD-MIM comparison fairly comprehensively. The use of a dimensionless variable does help in thinking about the key timescales, but it would also be good to see this comparison done with the raw parameters. Omitted from their analysis was analysis of goodness-of-fit, error space and noise.

6.6 MIM and DP matching

6.6.1 Comments on MIM and DP matching

The literature on how to relate the dimensionless MIM transfer coefficient, κ', to geometry is extensive[43]. The same problem has been posed in different contexts (for head as well as solute diffusion). For example these include groundwater (Zimmerman et al., 1993); soils (Gerke and van Genutchen, 1993); petroleum reservoirs (Kazemi et al., 1992); and CO_2 or CH_4 movement in coal (Boyer et al., 1982; Pavone and Schwerer, 1983). The comparison also features in the chemical engineering and chromatography literatures.

However, except in a few special examples (such as a reactor bed, particularly well-sorted soils or uniformly fractured rock), the assumption of similar block shapes and diffusion coefficients over a wide area in heterogeneous geological, soil or waste media seems highly unlikely, so this discussion is often rather hypothetical. But, the discussion is not always presented as such. The language of comparison of the DP versus MIM model can infer that the DP model is 'reality' and the MIM model is just an approximation to this. When applied to real heterogeneous media, the likelihood that *either* model is close to reality is slim.

A particular focus of the literature comparing the models has been consideration of the difference in the MIM and DP models at early time, given a fixed κ' (which is conventionally set to the asymptotically-matched value, say $\kappa' = \pi^2 / 4$ for a block). The MIM model is then

[43] The literature comparing the two models extends to other forms of comparison. These include:
- •Smoothly varying periodic functions (Mathias, 2005; Wright, 2002; Zimmerman et al., 1993).
- •Early-time (Mathias, 2005).
- •Pulsed input C(t) (Griffioen, 1998).
- •Interrupted flow q(t) (Brusseau et al., 1997; Harvey et al., 1994, Reedy et al., 1996; Wright, 2002).
- •Arbitrary input of C(t): requiring solution to convolution integral (Bibby, 1981; Carrera et al., 1998).
- •Comparsion of the MIM model against data (Griffioen et al., 1998; Nkedi-Kizza et al., 1983).

thought to be increasingly 'inaccurate' at earlier times (Mathias, 2005). The language of "correct" versus "inaccurate" is reasonable if there is a known regular geometry. Alternatively "correct" could be a shorthand for "assumed" and "incorrect" is just a description of the relative mismatch. It is not always clear in some cases whether the literal or shorthand interpretations are being used. Convincing field datasets are needed to support the first interpretation.

Asymptotic matching is not the only possibility and holds no theoretical superiority over other forms of comparison, which may be preferable in different circumstances. For example Parker and Valocci (1986) suggested an AD-MIM & AD-DP match by comparison of second centred moments and Young and Ball (1997) used this to compare a cylindrical tube to the MIM model.

There is a physical relationship between the MIM and the DP model which can only be made at the late-time point at which the concentration distribution within the DP block appears almost linear. The gradient at the block's edge which governs the flux into or out of the block is now approximately the same as the gradient throughout the block. The average concentration will also be proportional to the gradient. So, the DP model will behave very similarly to the MIM model, with a slight (but reducing) error due to the condition that the concentration gradient must be zero at the centre of the block. For earlier times this condition is clearly not held, so the two models will match less well.

The MIM model can be pictured as conceptualising a uniform ('fully-mixed') spatial distribution of solute within the block (e.g. Harvey et al., 1994; Wright, 2002). Moench (1984) noted that if a 'fracture skin' was modelled around a block, the spatial variation within the block is much reduced. Alternatively the MIM concept can be used without assuming this uniform distribution, as it potentially relates the mobile concentration C_m and the *average* immobile concentration $\langle C_{im} \rangle$, without need to invoke a uniform concentration.

In the instance of an asymptotic match, a practical issue that is seldom addressed is establishing the time after which the two models are *sufficiently* matched. Kazemi (1969) presented an empirical condition for the time at which reasonable approximation occurs, but this requires revisiting more thoroughly. It would be useful to be able to state that: "after X much time the difference in prediction of the two models is below Y percent".

Dykhuizen (1990) started with the assumption that the fundamental mode (described by the first eigenvalue) will dominate in the long-time. He stated this assumption in terms of a spatial concentration and arrives at the same result as if the simple time comparison for the average

concentration had been made (i.e that $\kappa' = 3$ in order to match a DP(slab) model with a MIM model).[44]

Various authors have considered matching the early time, for example Zimmerman et al. (1993). The comparison of the expression for exchange from an infinite matrix gives $\kappa' = 1/\sqrt{\pi T}$. So for an exact match κ' needs to be time-dependent and will tend to infinity at $T = 0$.

A final point about matching (Barker, personal communication) is that the particular problem in hand will contain particular initial and boundary conditions leading to a specific transport solution. The solution may be a 'functional' of the BGF. For example, for diffusion out of a block into an immobile zone with zero flow is described by $\overline{C}_m = \dfrac{C_m^0 + \sigma C_{im}^0}{s\left[1+\sigma \mathrm{B}\left(\sqrt{st_{cb}}\right)\right]}$. The match point for this functional will not in general be the same as for matching $\mathrm{B}\left(\sqrt{st_{cb}}\right)$.

6.6.2 MIM and DP BGF matching

6.6.2.1 Late-time matching MIM and DP BGF

In common with the comparison for the other processes, the late-time gradients can be compared for an impulse input.

Haggerty et al. (2000) showed how the late-time BTC concentration could be written in terms of a memory function, here re-expressed as the BGF. So:

$$C_m(T) \approx -t_a \left(m_0 / V_m\right) \sigma \frac{d\mathrm{B_t}(T)}{dT} \tag{6.12}$$

where the immobile concentration $C_{im}(T)$ is $\mathrm{B_t}(T)$.

Late-time gradient of the concentration eluted from a block

Taking the DP(slab) as an example, the late time concentration is $\mathrm{B_t}(T) \sim 2e^{-A_0 T}$, where $A_0 = \pi^2/4$. This gives $\dfrac{d}{dT}\ln\left[\mathrm{B_t}(T)\right] \sim -A_0 = -\pi^2/4$ for late time.

For the MIM model $\mathrm{B_t}(T) = \kappa' e^{-\kappa' T}$ and $\dfrac{d}{dT}\ln\left[\mathrm{B_t}(T)\right] = -\kappa'$ (i.e. a log-lin gradient of $-\kappa'$ for all time).

[44] Also see Section 6.6.2.2.

For the infinite block $B_t(T) \sim 1/\sqrt{\pi T}$ and $\frac{d}{dT}\ln[B_t(T)] = -1/(2T)$ (i.e. a log-log gradient of $-1/2$).

Late-time gradient of the BTC

Taking the DP(slab) as an example, the late time concentration is:

$$C_m(T) \sim 2(m_0/V_m)t_a\sigma A_0 e^{-A_0 T} \tag{6.13}$$

giving $\frac{d}{dT}\ln[C_m(T)] \sim -A_0 = -\pi^2/4$, i.e. a log-lin gradient of $-\pi^2/4$ for late time.

For the MIM model, the late time concentration is:

$$C_m(T) \sim (m_0/V_m)t_a\sigma\kappa'^2 \exp[-\kappa T] \tag{6.14}$$

giving $\frac{d}{dT}\ln[C_m(T)] \sim -\kappa'$ (i.e. a log-lin gradient of $-\kappa$ for late time).

For an infinite block, the late time concentration is:

$$C_m(T) \sim 2(m_0/V_m)t_a\sigma/\sqrt{4\pi T^3} \tag{6.15}$$

giving $\frac{d}{dT}\ln[B_t(T)] = -3/(2T)$ (i.e. a log-log gradient of $-3/2$).

For BGFs ($B_t(T)$) that can be expressed as sums of exponentials, the long time log-lin gradient of the block impulse response (the BGF) will be identical to the log-lin gradient of the impulse response of the whole system (the BTC). The exception is the instance of a log-log linear 'infinite' block. Note, too that the infinite block BGF gradient is -1/2 whereas the BTC gradient is -3/2.

Table 5.6 in Section 5.3.2 shows the late time gradients for the DP and MIM models. The log-lin gradient matching can then take place for a slab $\kappa' = -\pi^2/4$ and for a sphere, $\kappa' = -\pi^2$. This is identical to the asymptotic match achieved by matching the first eigenvalues.

The late-time match is not often expressed in this way, but instead is compared using graphs of $C_{im}(T)$ or $\Gamma(T)$ asymptotically.

The asymptotic equivalence of general functions $f(t)$ and $g(t)$ can be written, using the final value theorem (Barker, personal communication):

$$\lim_{t\to\infty}\frac{\mathrm{f}(t)}{\mathrm{g}(t)}=\frac{\lim_{t\to\infty}\mathrm{f}(t)}{\lim_{t\to\infty}\mathrm{g}(t)}=\frac{\lim_{s\to 0+}s\bar{\mathrm{f}}(s)}{\lim_{s\to 0+}s\bar{\mathrm{g}}(s)}=\lim_{s\to 0+}\frac{s\bar{\mathrm{f}}(s)}{s\bar{\mathrm{g}}(s)}=1 \tag{6.15}$$

Here $\bar{\mathrm{f}}$ is replaced by the DP formulation and $\bar{\mathrm{g}}$ by the MIM formulation, giving $\lim_{s\to 0+}\frac{\tanh\left(\sqrt{S}\right)/\sqrt{S}}{\kappa'/(S+\kappa')}=1$. This is true regardless of the choice of κ'. There is therefore no single 'correct' choice of κ' to give an identical final value. This is just a restatement of the earlier observation that BGFs tend to 1 as t tends to infinity. Mathematically there is no reason why κ' could not be negative, or complex, although it would lose its physical interpretation. What is of more interest however, is how chose the functions are as they approach their final values. It is not immediately clear what criteria to apply to evaluate what constitutes a good asymptotic match. The most obvious are to ensure that (a) the final value is correctly reached and (b) the difference between the functions monotonically decreases (to a vanishingly small difference as $t\to\infty$).

Now, consider the comparison between the late time DP concentration (for a slab block geometry) with the MIM model. Note that throughout this section the comparisons are illustrated for DP(slabs) versus MIM. The same method could be readily applied to any geometry. The asymptotic behaviour of the DP impulse response in concentration is given by using $\lim_{T\to\infty}\left(2\sum_{n=0}^{\infty}e^{-A_nT}\right)=2e^{-A_0T}$ and the MIM impulse response is given by $\kappa' e^{-\kappa' T}$.

So clearly it is impossible to equate the two functions for all time. Instead one way to match them is to keep the ratio between them constant. This is achieved by setting $\kappa'=A_0=\pi^2/4$, in which case:

$$\begin{aligned}\text{DP(slab)}&: \ 2e^{-A_0T}\\ \text{MIM}&: \ A_0e^{-A_0T}\end{aligned} \tag{6.17}$$

Note that this is equivalent to matching the exchange flux, q_e and differentiating Equations (6.17) w.r.t T, only introduces a factor of $-A_0$ to both models. Although the exponential powers are matched by this choice of $\kappa'=A_0$, there is still a relative error of $2/A_0$, whatever the choice of κ' (as originally pointed out by Barker (1985b)). However, this choice gives the most obvious consistent match over all long t, never getting worse than a factor of $8/\pi^2$; the error diminishing in magnitude with t. So it would not be right to claim that this first eigenvalue match is 'exact' or 'correct', yet it is still useful.

If a flushing system was considered, a log-C axis shows the models converging over long time (Figure 6.8). In the presence of noise and with limited measurement accuracy, there will be a time after which the two models are indistinguishable. If the noise, σ_{C/C_0}, is known, then the time after which the two models are within a given noise tolerance could be read off from a plot of $(C_{\mathrm{MIM}} - C_{\mathrm{DP}})/C_{\mathrm{MIM}}$ against t. For some applications when order of magnitude reductions in Cs are important and the concentration can be determined accurately over this range, the first eigenvalue fit would be useful. However, for intermediate time and/or when there is limited accuracy in the C measurement, the lowest eigenvalue is not necessarily optimum.

Figure 6.8 shows the concentration in a block that is flushed. The ratio of solutions for all time is the factor $\pi^2/8 = 1.2337$.

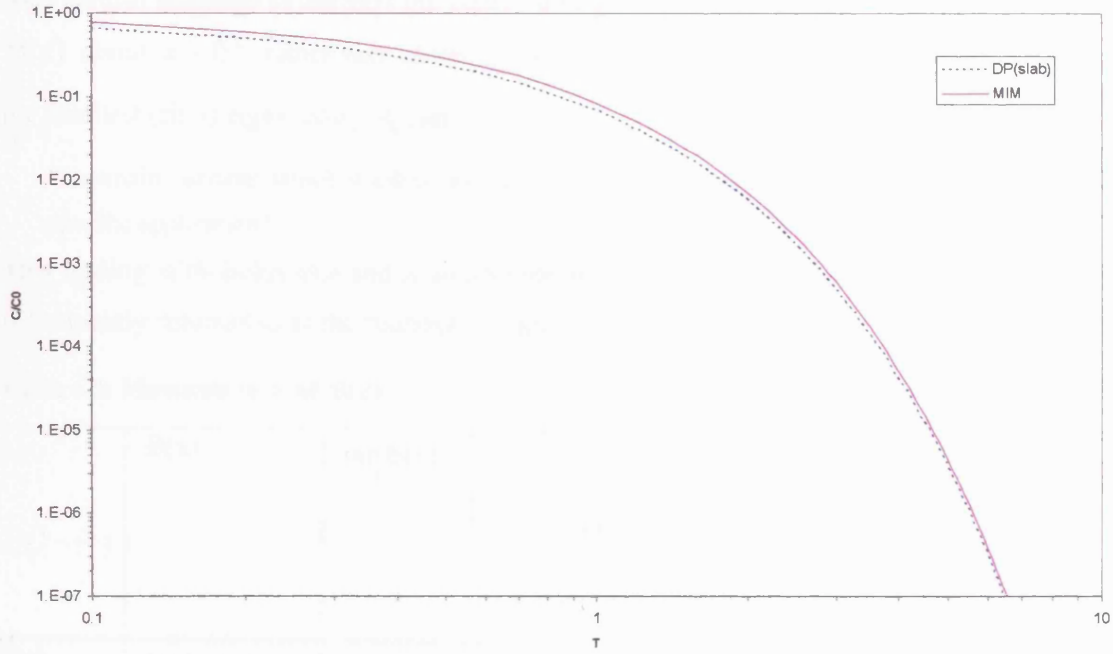


Figure 6.8: Block flushing response (i.e. initial C_m^0 and C_{im}^0 and input of $C_m(0,t) = 0$), with $\kappa' = \pi^2/4$, $t_a/t_{cb} = 10$, $\sigma = 1$, and $T = t/t_{cb}$.

6.6.2.2 Moment matching MIM and DP BGF

Section 5.3.3.2 demonstrated that as long as the mean immobile zone residence time is matched, then the zeroth, first and second moments of the BTC are matched (Cunningham and Roberts, 1998; Haggerty et al., 2000). The mean residence time is calculated by Haggerty et al. (2000) as the harmonic mean of the density function of MIM coefficients. However, it is just as easy to

find the mean by the method of moments without invoking an equivalent infinite sum of first order coefficients. Haggerty et al. (2000) described this as the 'best' choice of effective rate constant, in contrast to asymptotic being best. The choice clearly *depends on* the aims (although it must be said that the 'best' asymptotic value may be impossible to ascertain as it requires knowledge of the block sizes). Barker (1985b: Equation (44)), gave an expression for κ' to be matched:

$$\kappa' = -\frac{2}{\mathrm{B}''(x)\big|_{x=0}} \tag{6.18}$$

where $\mathrm{B}''(x)\big|_{x=0} = \frac{d^2}{dx^2}\mathrm{B}(x)\big|_{x=0}$ and $x = \sqrt{st_{cb}}$.

This resulted from taking a McLaurin expansion for $\mathrm{B}(x)$. Mathias and Zimmerman (2003) subsequently criticised this as being "incorrect", due to being it based on the "supposed small s – large t" relationship. Although the small s – large t is indeed erroneous, Barker (1985b) used careful language to describe the match, simply stating that, "it seems appropriate to expand $\mathrm{B}(x)$ about $x = 0$", rather than claiming any exactness. Barker (1985b) gave the solution of the smallest (first) eigenvalue, A_0, and commented on their relative merits saying:

> "it remains unclear which if either provides the best [κ'] value: presumably, this depends on the specific application".

This finding stills holds true and is worth reiterating in the context that the asymptotic solution is frequently referred to as the "correct" solution.

Table 6.9: Moments in x of $\mathrm{B}(x)$.

	B(x)	$\lim_{x\to 0}\mathrm{B}(x)$	$\mu_1^x{}' = -\frac{d}{dx}\mathrm{B}(x)\big\|_{x=0} = -\mathrm{B}'(x)\big\|_{x=0}$	$\mu_2^x{}' = \frac{d^2}{dx^2}\mathrm{B}(x)\big\|_{x=0} = \mathrm{B}''(x)\big\|_{x=0}$	$\mu_3^x{}' = -\frac{d^3}{dx^3}\mathrm{B}(x)\big\|_{x=0} = -\mathrm{B}'''(x)\big\|_{x=0}$	μ_2^x
Slab	$\tanh(x)/x$	1	0	$-\frac{2}{3}$	0	$-\frac{2}{3}$
Cylinder	$\frac{\mathrm{I}_1(2x)}{x\mathrm{I}_0(2x)}$	1	0	-1	0	-1
Sphere	$\coth(3x)/x - 1/(3x^2)$	1	0	$-\frac{6}{5}$	0	$-\frac{6}{5}$
MIM	$\frac{\kappa'}{x^2+\kappa'}$	1	0	$-\frac{2}{\kappa'}$	0	$-\frac{2}{\kappa'}$

From examination of Table 6.9 it is clear that a restatement of Barker's McLaurin expansion match is to equate the second raw moments $\mathrm{B}''(x)\big|_{x=0}^{DP} = \mathrm{B}''(x)\big|_{x=0}^{MIM} = \mu_2^{x\,MIM}$ thereby yielding Equation (6.18).

Applying a change in variables $st_{cb} = x^2$ by the chain rule to examine the condition for κ', in terms of s, gives:

$$\frac{t_{cb}}{\hat{\mathrm{B}}'(0,t_{cb})} = \frac{2}{\mathrm{B}''(x)\big|_{x=0}} \tag{6.19}$$

Recall that the first raw moment of the BGF is defined as $\hat{\mathrm{B}}'(0,t_{cb}) = \frac{d}{ds} B\left(\sqrt{st_{cb}}\right)\Big|_{s=0} = -\mu'_1$.

Therefore, $\kappa' = -\frac{2}{\mathrm{B}''(x)\big|_{x=0}} = -\frac{t_{cb}}{\hat{\mathrm{B}}'(0,t_{cb})} = \frac{t_{cb}}{\mu'_1}$. Substituting $\mu'_1 = \theta_{im}/\kappa$ for the MIM BGF gives $\kappa' = \frac{t_{cb}}{\mu'_1} = \frac{\kappa t_{cb}}{\theta_{im}}$. It is therefore clear that this match point is the same as matching the first moments of the BGF. So for the slab geometry, this first moment match gives $\kappa' = 3$, for the cylinder $\kappa' = 8$ and so on. Note that the same result can be achieved by comparing the McLaurin series for MIM and diffusive terms for both $\mathrm{B}(x)$ and $\mathrm{B}\left(\sqrt{st_{cb}}\right)$. This is a new way to interpret the match found by Barker (1985b) and Gerke and van Genutchen (1993).

If spread of tracer leaving a block is most important, then the second temporal moment could be matched (i.e. $\kappa' = \sqrt{-\frac{2t_{cb}^2}{\hat{\mathrm{B}}''(0,t_{cb})}}$). This second moment match is compared in Table 6.10 with the match of the first temporal moment.

The third moments would match if the $\hat{\mathrm{B}}''(0,t_{cb})$ s were also equal, but this cannot happen at the same time as $\hat{\mathrm{B}}'(0,t_{cb})$ being equal. However, if the second moment is matched it can be seen that the error in the third moment is small as long as $\gamma = t_a/t_{cb}$ is sufficiently large. Plotting this reveals a very good fit when κ'=3 and $\gamma = 10$ (Figure 6.10).

Table 6.10: Comparison of exchange coefficient choice for asymptotic and moment matches of the DP and MIM models.

Geom.	Asymptotic match	First raw moment match $\kappa' = -\frac{t_{cb}}{\hat{B}'(0,t_{cb})}$	Second raw moment match $\kappa' = \sqrt{-\frac{2t_{cb}^2}{\hat{B}''(0,t_{cb})}}$
Slab	$\pi^2/4 =$ 2.4674...	3	$\sqrt{\frac{15}{2}}$=2.7386...
Cylinder	1.45	2	$\sqrt{3}=1.7321...$
Sphere	$\pi^2/9 =$ 1.0966...	$\frac{5}{3}$=1.6666...	$\sqrt{\frac{35}{18}}$=1.3944...

6.6.2.2.1 Least squares matching of MIM and DP BGFs

There is no record of this in the literature, although it would seem a perfectly reasonable way to compare the BGFs. Subjective decisions would have to be made on the spacing of data and the weighting of the least-squares. This has not been performed here, since the major focus is on matching BTCs.

6.6.3 MIM and DP BTC matching

6.6.3.1 Late-time matching MIM and DP BTC

As long as Pe is not very small this comparison is identical to the AD-MIM & AD-DP comparison (see Section 6.5.2).

6.6.3.2 Moment matching MIM and DP BTC

Having performed the AD-MIM against AD-DP comparison the matching of moments for MIM and DP is now trivially achieved by setting $t_D = 0$ to cross out all the AD terms in the former more general comparison.

Table 6.11: Comparison of MIM and DP BTC moments.

Geometry	μ_2	μ_3
Slab	$\frac{2}{3}t_a\sigma t_{cb}$	$\frac{4}{5}t_a t_{cb}^2\sigma$
Sphere	$\frac{2}{15}t_a\sigma t_{cb}$	$\frac{4}{105}t_a t_{cb}^2\sigma$
MIM	$\frac{2}{\kappa}t_a\sigma\theta_{im}$	$\frac{6}{\kappa^2}t_a\sigma\theta_{im}^2$

For the second moment match of MIM and DP(sphere), $\frac{t_{cb}}{15} = \frac{\theta_{im}}{\kappa}$, or $t_{cb} = 15\theta_{im}/\kappa$. For the third moment, $\frac{2}{315}t_{cb}^2 = \frac{\theta_{im}^2}{\kappa^2}$, or $t_{cb} = \theta_{im}\sqrt{315/2}/\kappa \approx 12.55\theta_{im}/\kappa$. Therefore the second and third moments in general will not both match.

6.6.3.3 Least squares matching of BTCs for MIM and DP models

This issue has been covered extensively in the literature. Griffioen et al. (1998) gave a detailed analysis of 20 sets of experimental data. Nkedi-Kizza et al. (1983) analysed a series of soil experiments.

The observation of κ scaling with size of experiment is made by Haggerty (1999) who plots κ and $1/t_{cb}$ against the timescale of the experiment for 100 experiments. Because the diffusive data also shows a similar relationship to timescale as MIM did in Young and Ball (1995), it is suggestive that there are further issues at work beyond the explanation that an inappropriate model is being fitted. Liu et al. (2004) presented evidence of scale dependence of D_a (this is covered in more detail in Section 7.3.3). So the work of Young and Ball (1995) is useful, but is not definitive.

The same comparison can be made for the BTC, which is often the point of practical comparison between models. Consider the comparison with $\gamma = t_a/t_{cb} = 10$ and $\sigma = 1$ (Figure 6.9 and Figure 6.10).

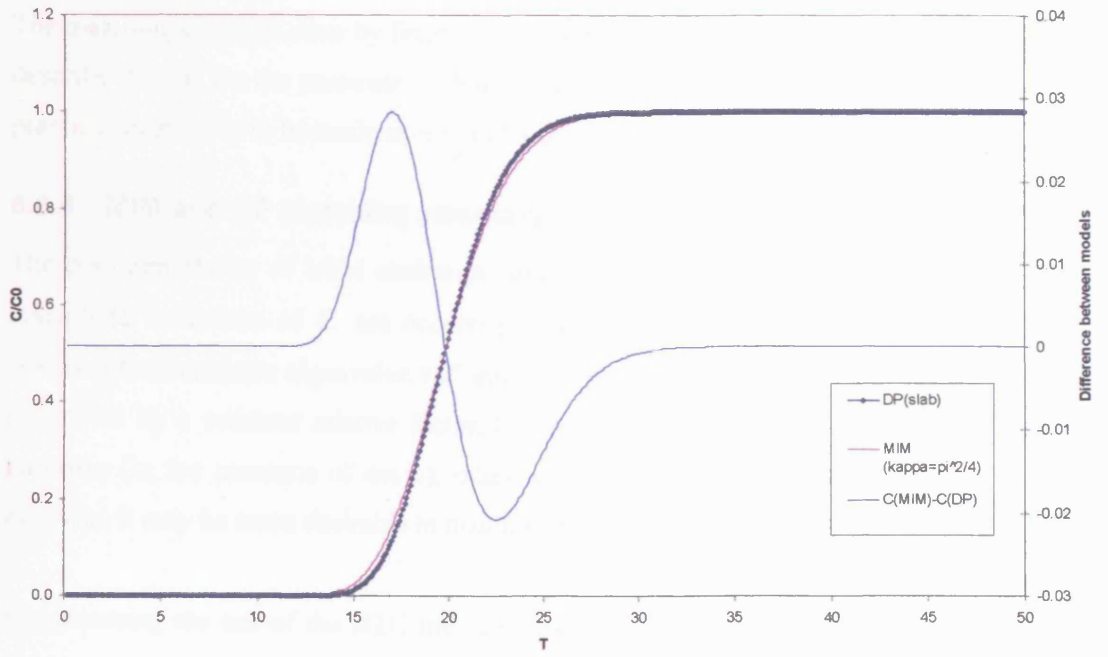


Figure 6.9: BTC step response for $\kappa' = \pi^2/4$, $\gamma = t_a/t_{cb} = 10$, $\sigma = 1$ and where $T = t/t_{cb}$.

The two curves are very close indeed. Lowering the exchange parameter to $\kappa' = 3$ gave a better visual fit. This is of course also the point at which we know the first moment for the BGF and hence the second centred moment for the BTC is matched.

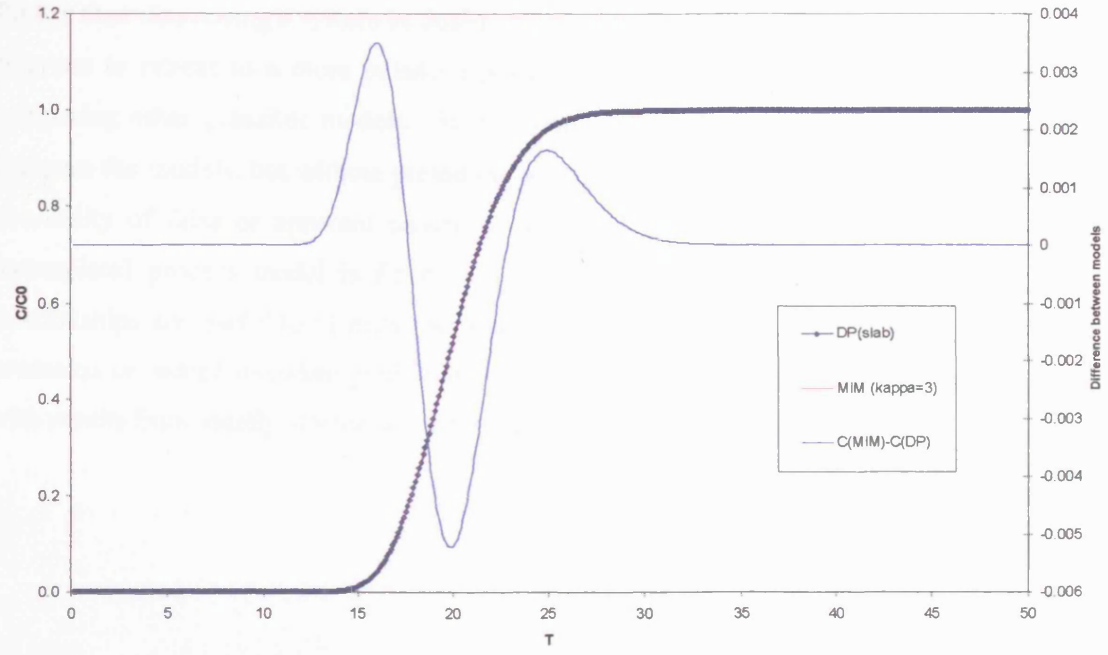


Figure 6.10: BTC step response for $\kappa' = 3$, $\gamma = t_a/t_{cb} = 10$, $\sigma = 1$

The matching could be done by fitting to synthetic data with noise added and by MC analysis to describe the CR for the parameters. Such an analysis (as future work) would potentially allow practical statements to be made about model equivalence given the data.

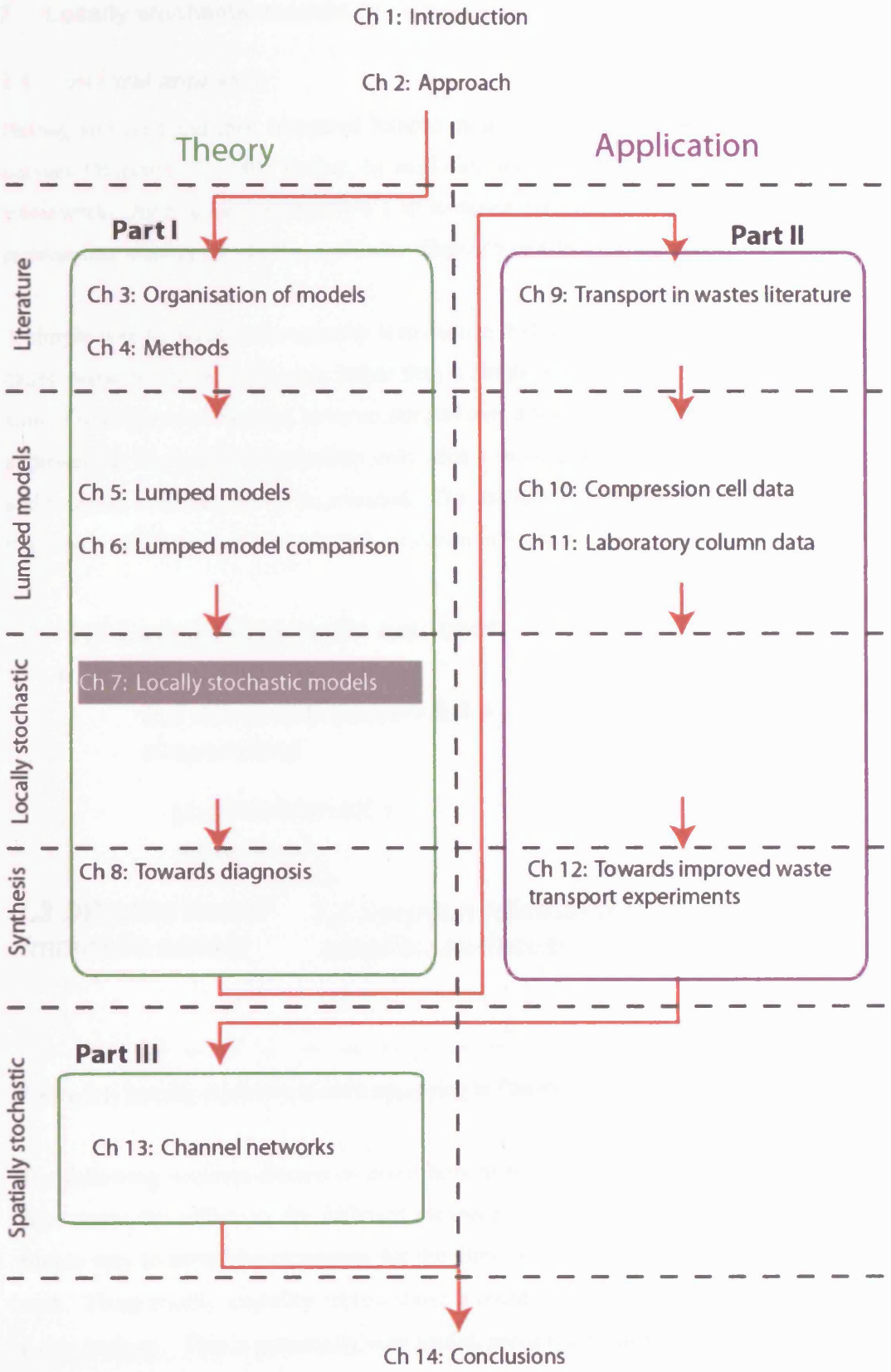
6.6.4 MIM and DP matching summary

The optimum choice of MIM exchange parameter, κ, depends upon application. If order of magnitude reductions of C are occurring and important and if the flow-geometry is known, selecting the minimum eigenvalue will approximately match the tail. Even so, it will always be inaccurate by a constant relative factor, but by a diminishing absolute amount. For all other purposes (in the presence of noise), other κ values may be more practical and useful. For example, it may be more desirable to match mid-time, perhaps the peak or moments.

For matching the tail of the BTC the same issues arise as for the BGF. For other times, the results are better than expected and the system is quite 'forgiving'. This is perhaps due to the fact that the first moment will always match, and the second will match if the first moment of the BGF is matched.

The match presented by Barker (1985a) and Gerke and van Genutchen (1993) is actually equivalent to matching the first moment. In the presence of noise, this may be a considerably better choice than the 'asymptotic' value.

Rather than diagnosing a system as dual-porosity without very good field data it would be more rigorous to retreat to a more balanced position of confronting the DP model with data and comparing other plausible models. In the event of 'equifinality' (Section 2.5) it is useful to compare the models, but without presuming which is 'correct'. A key issue that emerges is the possibility of false or apparent parameter scaling relationships that arise when a wrong (or incomplete) process model is fitted to a system dominated by a different process. Such relationships are useful to identify, because they could account for mistaken deductions about processes or indeed mistaken predictions. The results here are summarised in Chapter 8 along with results from locally stochastic models (see Chapter 7).


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

7 Locally stochastic models

7.1 *Integral approach*

Having analysed and then compared lumped models and coupled lumped models, this chapter pursues Objective 1 a little further, by exploring the next level of heterogeneity in the model framework. As a reminder, objective 1 is: *to develop a comparative framework of a sub-set of appropriate models for clearer and better diagnosis of simple processes.*

A simple way to model heterogeneity is to assume that there is a distribution of components that cause particular basic processes, rather than a single lumped process. This can be done by a sum of discrete contributions, or by an integral over a smooth distribution of contributions. This approach (at its simplest) potentially only adds a single parameter (the spread of a distribution) and requires a distribution to be assumed. The sections in this chapter correspond to corners of the locally stochastic process triangle, as shown in Figure 7.1.

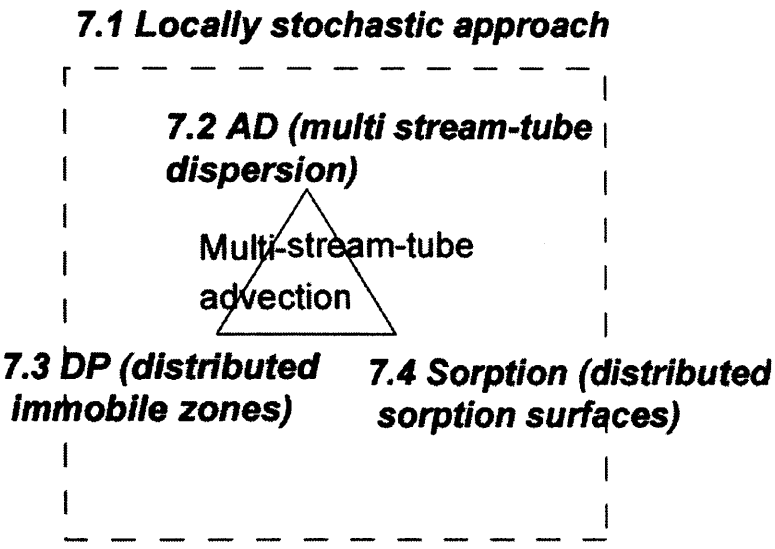


Figure 7.1: Locally stochastic models appearing in Chapter 7.

The following sections discuss in detail how to form such a heterogeneity representation and then assess the effect on the different measures of the BTC. A commonality is found in a simple way to model heterogeneity for the three basic processes, by summing components for each. These readily modelled representations could add a broad spectrum of possible models to a risk analysis. This is potentially very useful, particularly since conceptual uncertainty may be as important as parameter uncertainty.

The main contribution of this chapter is to examine locally stochastic heterogeneity in an organised way with a single coherent nomenclature. This is essentially an organised critical review of existing approaches. Within this review certain new results are also added. In

particular relationships between pdf of block sizes and block times are found in Section 7.34 and a generalised moment matching is derived in Section 7.3.8, which is based on the less general results of Cunningham and Roberts (1998).

7.2 *Locally-stochastic AD representation*

The first corner of the triangle in Figure 7.1 to be considered is AD. Instead of having a single lumped AD equation, it is possible to have a summation of parallel AD systems that exist at smaller scales (i.e. a multi stream-tube approach: see Figure 7.2). Doing this enables a wider range of systems and BTC phenomena to be modelled. Also, it may be that lumped dispersion fails to correctly capture the extent of local mixing and therefore affects reaction rates in the instance of modelling reactive transport (Cirpka and Kitanidis (2000a) made the point that it is molecular-scale mixing that affects the reaction rate).

The simplest of such models is the addition of weighted discrete pathways, giving $\bar{C} = \sum_{i=1}^{N} w_i C_i$. Such methods include the 'multi-dispersion' model of Maloszewski et al. (1992). In this model the pathways are equally weighted, there is a constant Pe and a particular t_A is allocated to each pathway.

Discrete models are useful if a relatively small number of distinct pathways are anticipated (perhaps identifiable channels) and the BTC is visibly multimodal.

More generally the concentration can be written as a continuum of weighted contributions over many different parameters,

$$\bar{C} = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} C(\eta_1 \ldots \eta_n) p(\eta_1 \ldots \eta_n) d\eta_1 \ldots d\eta_n \qquad (7.1)$$

Using the simplest possible process, the piston-flow model (Jury and Roth, 1990), gives so-called 'stochastic-convective models'. Adding the distances travelled for many parallel non-intersection piston flow models (each with random V) and applying the Central Limit Theorem (CLT) gives a spatial Gaussian distribution of distance travelled (Flühler et al., 1996).

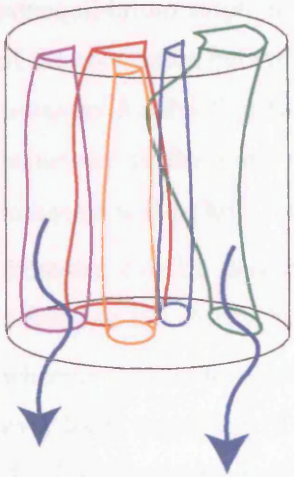


Figure 7.2: Multiple stream-tubes.

A more complex description applies when there is smaller scale mixing in the tubes. This is a physically plausible concept, as processes are observed that generate AD at channel or single-fracture scales (e.g. 'Taylor dispersion').

Cirpka and Kitanidis (2000) achieved this method by varying the arrival time, t_A, in a continuous version of the discrete model by Maloszewski et al. (1992). The multi-dimensional expression of Equation (7.1) is here only applied for t_A, which collapses to: $\hat{C} = \int_{-\infty}^{\infty} C(\eta_1) p(\eta_1)\, d\eta_1$, so $\hat{C} = \int_0^{\infty} C(t_A) p(t_A)\, dt_A$ (note the change of integration limits as t_A is not expected to be negative).

The semi-infinite impulse response is $C = (m_0/V)\frac{t_A}{\sqrt{\omega\pi t^3}} e^{\frac{-(t_A-t)^2}{\omega t}}$, where $\omega = 4\frac{t_A}{Pe}$. So the complete integral is:

$$\hat{C} = (m_0/V)\int_0^{\infty} \frac{t_A}{\sqrt{\omega\pi t^3}} \exp\left[-\frac{(t_A-t)^2}{\omega t}\right] p(t_A)\, dt_A \qquad (7.2)$$

where Pe is kept constant between the stream-tubes.

Toride and Leij (1996a&b) describe a methodology that is embodied in the CXTFIT code (van Genuchten, 1999). This is slightly more general than the above case in as much as a MIM exchange is also included (i.e. an AD-DP(MIM) model). They allow the integral to be over only two parameters, V and one of D, κ or K_d (where κ and K_d describe first-order

nonequilibrium sorption).[45] They select a bivariate lognormal distribution and explore variation of the correlation between the pairs of parameters. The lognormal distribution allows analytical solutions for the first two moments. As long as the mean of the distribution of a particular parameter is the same as the value allocated for a single stream-tube case, then the first moments will be kept the same. For the case of a bivariate distribution of V and D, the second moments can be decomposed into a MIM term, a D term and a term characterising the difference in velocity between the stream-tubes. The first two terms are linear in distance, z, whereas the last term is in z^2. So, the last term, the stochastic-advective term, will dominate over local variability of other parameters. In other words the 'locally stochastic' component dominates over the dispersion term in each stream tube.

The spatially integrated BTC reflects both the variance of the arrival times in each stream tube as well as the average of the variance of the local second central moments (i.e. within each tube), as described by Cirpka and Kitanidis (2000a).

Possible instances in which distinct preferred routes exist include the early 'fingering' stages of infiltration (Flühler et al., 1996), macropores in soils made up from burrows and root-moulds, karstic channels and horizontal flow along different strata and along parallel fractures. In an experiment the flow along the wall of the column might act as a distinctly different stream-tube to the rest of the flow pattern. Beyond these clear examples there probably exists a continuum of systems in which there is increasing lateral interaction, for which the non-interacting stream-tubes are an end-member.

There are a large number of ways of combining stream-tubes. Neretnieks (2002) considered in a 'multichannel' model a distribution of flow rates as well as channel volume, exploring the effect of how channel volume (parameterised as aperture) relates to flow rate and hence residence time. If there is little mixing within the stream-tubes (or transfer between them), this model gives a dispersion length that is proportional to distance (therefore predicting the scale effect).

For a particular situation it may be obvious which approach should be selected, but it may also be a relatively arbitrary choice. In the latter instance it is well-worth testing the sensitivity of the modelled outcome to the method of combination.

[45] This is a further way of representing heterogeneous sorption in contrast to the method given in Section 7.4. It is the same as found by Rabideau and Miller (1994) for horizontal flow through vertically stratified lognormally distributed K (i.e. the greatest spread occurs if K_d is negatively correlated with V (or K)).

7.2.1 Late-time result for AD

The late-time concentration for the locally-stochastic AD model is investigated for two simple distributions of t_A - a Gaussian and a lognormal distribution using a code called HET-STREAMS.F90 (see Appendix 1).

7.2.1.1 Normal distribution of t_A

The comparison between a single stream-tube and stream-tubes with t_A normally distributed is shown in Figure 7.3. Parameters are set to V =1, z =5 (hence the mean $\hat{t}_A = 5$), α =1 and SD =1. Note that the distribution is chosen such that the probability of negative t_A is insignificant. The mean time, $\hat{t}_A$, is kept the same for both. There is relatively little impact on the BTC due to the distribution of t_A and the late-time curves are effectively the same. It is only the early time that is affected. Note that for SD <1 the difference is insignificant (and not shown here).

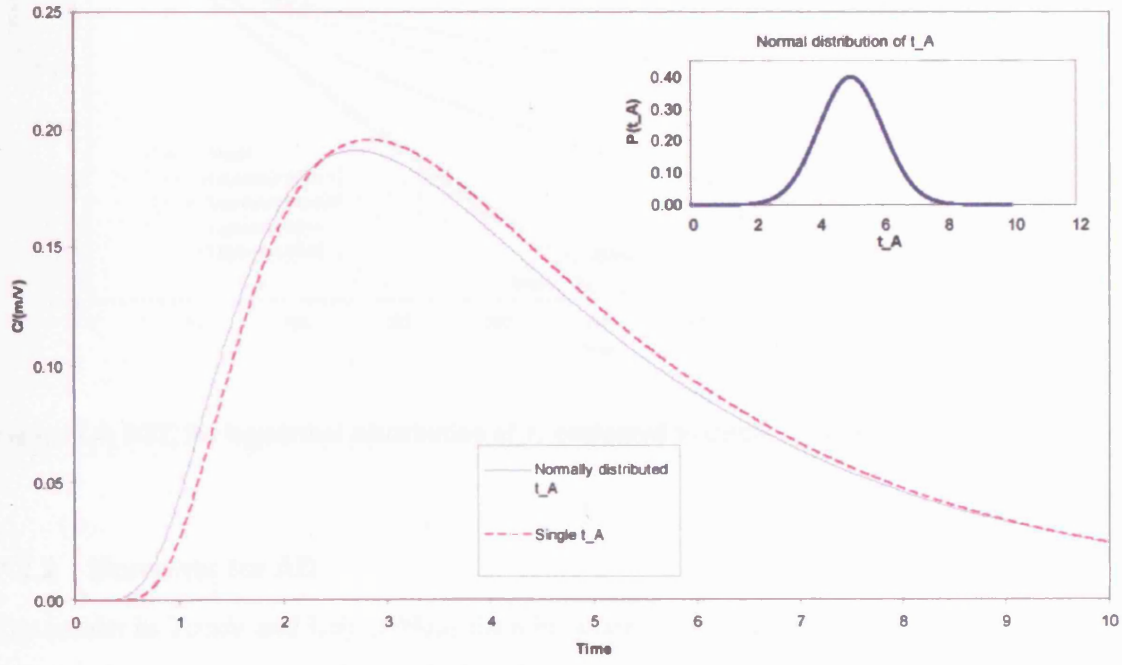


Figure 7.3: BTC for locally-stochastic AD with Normal distribution of t_A compared to a single t_A, on linear axes (t_A is distributed as $N(\hat{t}_A = 5, SD = 1)$).

7.2.1.2 Lognormal distribution of t_A

The comparison between a single streamtube and stream-tubes with t_A lognormally distributed is now simulated. Parameters are set to V=1, z=5 (hence mean $t_A^0 = 5$) and α=1. For the lognormal distribution (i.e. $P(t_A) = \frac{1}{t_A SD\sqrt{2\pi}} \exp\left[-\frac{\left(\ln(t_A)-\hat{t}_A\right)^2}{2SD^2}\right]$), different values of SD are plotted. The lognormal mean time, $\hat{t}_A$, is calculated as $\hat{t}_A = \ln\left(t_A^0\right) - \frac{1}{2}SD^2$. Figure 7.4 shows that there is substantial impact on the BTC at late-time.

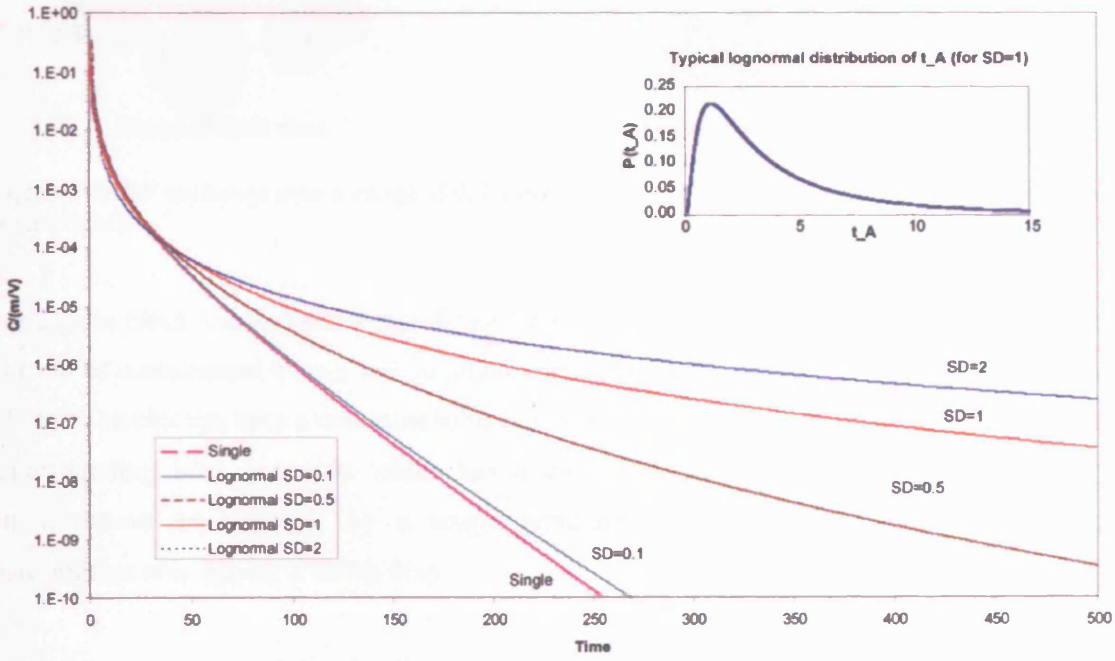


Figure 7.4: BTC for lognormal distribution of t_A compared to single t_A, on log-log axes.

7.2.2 Moments for AD

The results in Toride and Leij (1996a) for a bivariate lognormal are the only analytical results that have been found. Extending to further distributions would be a useful future investigation.

7.3 *Locally-stochastic Dual Porosity representation*

7.3.1 Introduction to locally-stochastic approaches to Dual Porosity

The central idea behind 'distributed diffusion' approaches is that there is no single representative scale over which mass-transfer interaction takes place. This idea is not by any means new. The concept can be traced back to the chemical engineering literature and the work

of Ruthven and Loughlin (1971). A burst of publications on the issue arose in the mid 1990's (see Haggerty and Gorelick (1995) and references therein). Rather than having a single homogenised 'average' parameter to model mass-transfer, a distribution of scales is modelled (Figure 7.5). Clearly some systems are regular (for example cleated coal, certain fractured rock systems and well-graded granular systems), but these may be extreme cases. To cover a wider range of materials a more general formulation is required.

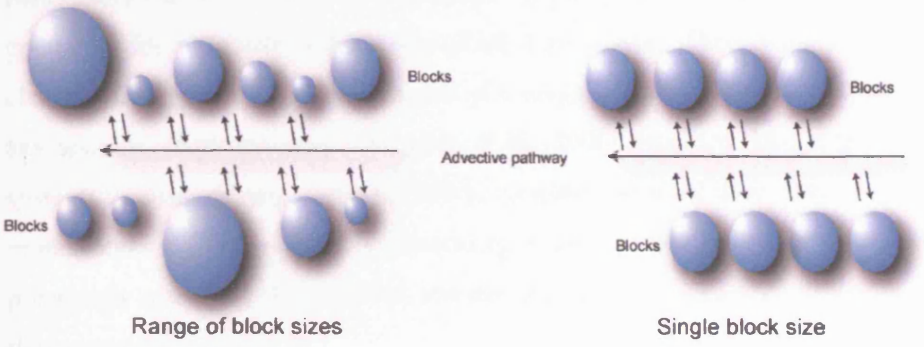


Figure 7.5: DP exchange over a range of diffusion scales (left), rather than for a single homogenised scale (right).

A range in block sizes could be significant for problems in which the time to flush the last few percent of contaminant from a soil, groundwater or landfill is critical. In such cases the fraction of larger blocks can have a dominant influence on late-time behaviour. It is also very useful for understanding how a system which has a highly variable block size might under some circumstances be captured by a single representative block size (and where potential inaccuracies may appear if doing this).

There are examples of the distributed-diffusion approach in a range of applications from granular soils to fractured crystalline rock. The wide applicability of all these approaches is likely to be due to the fact that they are highly flexible and probably capture an 'essence' in a simple way, which is still very general.

Simpler single representative scale DP models may become inadequate when confronted with datasets that have orders of magnitude reductions in concentration. In this instance, concentration is best plotted on a log-scale in order to see what happens in the tail. Cunningham and Roberts (1998) have shown that as long as the mean of the distribution of scales in the BGF is matched to the single scale, then the first two moments of the impulse response BTC will match. This semi-quantitatively explains why a system that is really made up of a distribution of scales is adequately modelled by a single scale for relatively small amounts of flushing. So, when concentration is viewed on a non log-scale and the flushing of the last few percent is not a

great priority, single-scale models can often be adequate, and the additional complexity of invoking multiple scales would seem unnecessary.

Heterogeneous sorption will also contribute a potentially indistinguishable additional heterogeneity of times (see Section 7.4). It is worth noting that there may be instances where there is insufficient knowledge to separate the contributions from diffusion, kinetics sorption and heterogeneous sorption, in which case a model may be fitted such as to find bulk apparent parameters that lump all the contributions. If the system is indistinguishably jumbled, there is a good case for summing distributions of MIM processes. These have the advantage that they are characterising the range of timescales of a range of processes and are less obviously related to any specific single process. Haggerty et al. (2000) confined their approach to representing all systems as sums of many parallel MIM processes. In some cases this can add complexity to the representation (for example, if simulating a single diffusion scale into a block). However, it potentially also adds the practical advantage that a fast computation can be used, which avoids the numerical inverse LT.

7.3.2 Mathematics of distributed DP

Consider a DP equation that is modified by replacing the BGF for a single diffusion time, $\mathrm{B}\left(\sqrt{st_{cb0}}\right)$, by the integral

$$\int_0^\infty p(t_{cb})\left(\mathrm{B}\sqrt{st_{cb}}\right)dt_{cb} \tag{7.3}$$

For distributed diffusion the block response function $\mathrm{B_t}\left(t,t_{cb}\right)$ is replaced by a parallel and simultaneously acting sum of blocks, $\mathrm{B_t}(t)=\int_0^\infty P(t_{cb})\mathrm{B_t}\left(t,t_{cb}\right)dt_{cb}$. Similarly, the LT is just $\mathrm{B}(s)=\int_0^\infty p(t_{cb})\mathrm{B}\left(s,t_{cb}\right)dt_{cb}$.

For example, for the DP-only model $\bar{C}_f=\left(m/V_m\right)\exp\left[-t_a s\left(1+\sigma\mathrm{B}\left(\sqrt{st_{cb}}\right)\right)\right]$ is replaced by

$$\bar{C}_f=\left(m/V_m\right)\exp\left[-t_a s\left(1+\sigma\int_0^\infty P(t_{cb})\mathrm{B}\left(\sqrt{st_{cb}}\right)dt_{cb}\right)\right] \tag{7.4}$$

A special case is the instance of a single diffusion time, $p(t_{cb})=\delta(t_{cb}-t_{cb0})$. In this case $\mathrm{B}(\sqrt{st_{cb0}})$ is recovered.

In general, this integral is difficult to determine analytically. For computational purposes $\mathrm{B}(s)$ can be made into a truncated infinite sum, $\mathrm{B}(s) = \sum_{i=1}^{M} p(t_{cbi})\mathrm{B}\left(\sqrt{st_{cbi}}\right)$, where M is a large finite number. The remaining $p(t_{cb})$ for $i > M$ is lumped into $p(t_{cbM})$ and allocated time t_{cbM}. This is essentially saying that the higher terms equilibrate so quickly that they are not worth allocating separate exchange parameters. Haggerty and Gorelick (1995: 2389) called this 'truncation control'. The relative time of the experiment, t_a, and the smallest block times will affect the calibration of θ_m, since the fraction which comes quickly into equilibrium with the mobile zone effectively becomes part of the mobile zone. This observation may be applicable to experiments which show calibrated changes in σ with different process conditions (for example some of the findings of Griffioen (1998) may be influenced by such effects).

7.3.3 In support of the DP heterogeneity method

A major objection to this method is that the distribution might be very hard to constrain by measurement and may be non-unique by calibration. One response to this is that it is useful to seek extreme but plausible distributions that give good fits to datasets, but potentially diverging predictions for extrapolation to long-time or for changes in process conditions. Retreating to a single-time model apparently without a distribution of time does not avoid assuming a distribution (just that it is now $\delta(t - t_{cb})$)[46]. Furthermore, for the simplest monomodal distributions only two additional assumptions need be made - the type of distribution and a further parameter characterising its spread.

The inability of many current datasets to provide any constraint for distributed models does not necessarily invalidate the approach. Instead, by suggesting a model of this form, a data requirement challenge is being set.

One further reason why this approach may be appropriate is that even with a very regular physical geometry there is a high probability of preferred pathways creating an irregular flow geometry. In other words, physical measurement of geometry might not provide adequate constraint; what is really needed is full three dimensional characterisation of flow paths, which is not currently available.

[46] Analogously, finding the 'unweighted' least squares can be thought of as finding the weighted least squares, but with all weights assumed to be equal to one. Unless a rationale is given for doing this, choosing any other distribution of weights might seem to be equally arbitrary.

The approach is simpler than spatial stochastic representations and is potentially an attractive addition to simple risk models of contaminant flushing. In reality there will be spatial variation and this will probably affect both spatial and temporal measures of transport. However, the data requirement to correctly describe the spatial pattern is even more demanding than describing a block size distribution.

Haggerty (1999) noted an observation-time dependence of rate coefficients plotted for a large range of scale of experiment. This is possibly explained by calibration of a single t_{cb} in a system that is truly multi-rate and may give rise to a t_{cb} that is a function of the experimental time. Haggerty (1999) acknowledged the work of others (Griffioen et al., 1998; Li et al., 1994; Young and Ball, 1995) who found that V_m and 'contact time' could account for such a relationship.

In the context of fractured rock, Neretnieks (2002) and Shapiro (2001) noticed that the diffusion coefficient for field experiments needed to be considerably larger than it did in the laboratory. Liu et al. (2004) presented a graph of distance against diffusion coefficient (see Figure 7.6) for a number of different datasets, achieving a result analogous to the 'scale effect' graph shown by Gelhar (1993) for AD. Their preliminary explanation for this is that solute travel paths within a fracture network are fractals.

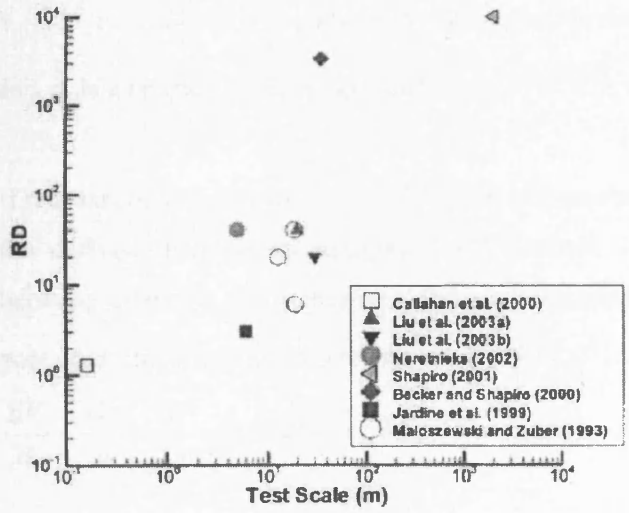


Figure 7.6: RD against test scale from Liu et al. (2004), where RD is the ratio of calibrated effective matrix diffusion coefficient to the measured value in the laboratory.

7.3.4 Choice of parameter to distribute

There is a choice of which parameter to distribute. Some different choices that appear in the literature are shown in Table 7.1.

Table 7.1: Different choices of distributed properties

Property	Author(s)
MIM times or transfer coefficients	Haggerty and Gorelick (1995), Sardin et al. (1991)
Block dimension	Valocci (1990), Rameson (1985)
Block volume	Barker (1985b)
Aggregate weight (probably proportion to volume)	Sardin et al. (1991)
Block times	Cunningham and Roberts (1998)[47]

These distributions are applied by different authors through different underlying block shapes (usually simple slabs, cylinders or spheres). Even more generally, some authors have considered the effect of a distribution of block shapes at the same time (see Barker, 1985b; Zimmerman and Bodvarsson, 1990).

Which of these distributions of properties is most appropriate may depend on the nature of the system and the data available. Haggerty and Gorelick (1998) made a case for distributing diffusion times independently of any distributions of characteristic sizes. The assumption is that diffusion takes place along many separate linear paths within individual 'grains'.[48] The alternative model is that the diffusion times are intimately related to the block sizes and there will be a conversion from volumetric distribution to times.

A literature search revealed no descriptions of the relationships between the distributions $\mathrm{f}_t(t), \mathrm{f}_a(a)$ and $\mathrm{f}_V(V)$, where V (in contrast to the rest of the thesis) is used to mean volume and a is a characteristic length scale and $t = a^2 / D_a$.

If it is assumed that there is pure diffusion into precise shapes, then the pdfs can be related using the derived distribution (Benjamin and Cornell, 1970; Gelhar, 1993) allowing conversion between different representations and allowing relation to different types of dataset. Firstly, note that the parameters are related by $V \propto a^3$ and $t \propto a^2$ (i.e. $V^{1/3} \propto a \propto t^{1/2}$), so that $\frac{dV}{3V} = \frac{da}{a} = \frac{dt}{2t}$ and therefore $3V\mathrm{f}_V(V) = a\mathrm{f}_a(a) = 2t\mathrm{f}_t(t)$.

Therefore the pdfs can be related by $\mathrm{f}_t(t) = \frac{3V\mathrm{f}_V(V)}{2t} = \frac{a\mathrm{f}_a(a)}{2t}$ and $\mathrm{f}_a(a) = \frac{2a}{D_a}\mathrm{f}_t\left(\frac{a^2}{D_a}\right)$.

Assuming spherical geometry and letting $a = r/3$, then gives:

[47] Cunningham also simultaneously introduced a linear sorption term and therefore a retardation coefficient.

[48] 'Grains' are used in the context of soil particles. However, the argument is potentially generalised to other block types.

$$f_V(V) = \frac{4}{D_a \pi^{2/3} V^{1/3} 6^{7/3}} f_t\left(\frac{\left(\frac{V}{36\pi}\right)^{2/3}}{D_a}\right).$$

A special result comes about for the lognormal distribution. If any one of a, V or t is lognormally distributed then so too will the others be.

A useful reason for being able to change variables is that there may be a known physical upper bound for one. For example, in the case of a column test, the largest block dimension is likely to be the radius of the cell.

7.3.5 Equivalent single BGF

All the distributed diffusion representations in the literature are fairly similar. However, the derivation and representation by Barker (1985b) is cast rather differently. The approach is based on defining a harmonic 'average' dimension $\bar{a}$ or time $\bar{t}_{cb}$.

There follows an elucidation of Barker's approach (Barker personal communication). The nomenclature has been slightly changed by this author to make the harmonic average values clear as 'averages' and to emphasise the fact that the total integrated or summed BGF is not really a single BGF but is still a sum (therefore replace a by $\bar{a}$ and $B(x)$ by $B_\Sigma(x)$ or $B_\int(x)$). The idea behind this nomenclature is to keep the complete mass transport equation simple. So instead of $\bar{C}_f = (m/V_m)\exp\left[-t_a s\left(1+\sigma\int_0^\infty p(t_{cb})B\left(\sqrt{st_{cb}}\right)dt_{cb}\right)\right]$, Barker expressed distributed diffusion as $\bar{C}_f = (m/V_m)\exp\left[-t_a s\left(1+\sigma B_\Sigma(x)\right)\right]$.

Discrete blocks are described by $B_\Sigma(x) = \sum_{i=1}^{N} p_i B_i\left(\frac{xa_i}{\bar{a}}\right)$, where $\frac{1}{\bar{a}} = \sum_{i=1}^{N}\frac{p_i}{a_i}$ and $a_i = V_i / A_i$.

For example, as applied to a bimodal distribution, with one group of block size a_1 occupying volumetric proportion P and the other of size a_2, $1-P$

$B(x) = PB\left(\frac{xa_1}{\bar{a}}\right) + (1-P)B\left(\frac{xa_2}{\bar{a}}\right)$, where $\frac{1}{a} = \frac{P}{a_1} + \frac{(1-P)}{a_2}$. Alternatively:

$B(\sqrt{st_{cb}}) = PB\left(\sqrt{st_{cb1}}\right) + (1-P)B\left(\sqrt{st_{cb2}}\right)$, where $\frac{1}{\sqrt{st_{cb}}} = \frac{P}{\sqrt{st_{cb1}}} + \frac{(1-P)}{\sqrt{st_{cb2}}}$.

This should not be confused with attempting to replace the integral $\int_0^\infty p(t_{cb})\mathrm{B}\left(\sqrt{st_{cb}}\right)dt_{cb}$ by the same BGF with a single time that gives in some way equivalent behaviour. In other words, Barker *does not state* that $\mathrm{B}(\sqrt{st_{cb_1}}) \equiv \int_0^\infty p(t_{cb})\mathrm{B}\left(\sqrt{st_{cb}}\right)dt_{cb}$.

Continuous blocks are described by $\mathrm{B}_\mathrm{J}(x) = \sum_{i=1}^{N_s}\int_0^\infty p_i(\alpha)\mathrm{B}_i\left(\frac{x\alpha}{\overline{a}}\right)d\alpha$, where $\frac{1}{\overline{a}} = \int_0^\infty \frac{p(\alpha)}{\alpha}d\alpha$

Barker further relates $P = \mathrm{f}_a(a)$ and $P = \mathrm{f}_t(t)$ (see Table 7.2).

Table 7.2: Relationship between different pdfs and B.

Parameter	$\mathrm{B}(x)$	**Where in B(x)** $\frac{1}{\overline{a}}$ **is given as:**	$\mathrm{B}_\mathrm{J}\left(\sqrt{s\overline{t}}\right)$
V	$\int \mathrm{f}_v(V)\mathrm{B}\left(\frac{xa}{\overline{a}}\right)dV$	$\int \frac{\mathrm{f}_v(V)}{\overline{a}}dV$	$\int \mathrm{f}_v(V)\mathrm{B}\left(\sqrt{st}\right)dV$
a	$\int \mathrm{f}_a(a)\mathrm{B}\left(\frac{xa}{\overline{a}}\right)da$	$\int \frac{\mathrm{f}_a(a)}{\overline{a}}da$	$\int \mathrm{f}_a(a)\mathrm{B}\left(\sqrt{st}\right)da$
t	$\int \mathrm{f}_t(t)\mathrm{B}\left(x\sqrt{\frac{t}{\overline{t}}}\right)dt$	$\int \frac{\mathrm{f}_t(t)}{\overline{a}}dt$	$\int \mathrm{f}_t(t)\mathrm{B}\left(\sqrt{st}\right)dt$

Barker (personal communication):

> "While $a/\overline{a} = \sqrt{t/\overline{t}}$, we can't relate the ratio $a/\overline{a}$ to $V/\overline{V}$, hence the 'asymmetry'. Define $\overline{t} = \frac{\overline{a}^2}{D_a}$, as $\overline{t}$ does not seem to have a meaning in its own right (as do $\overline{a}$ and $\overline{V}$ for the 'composite' block)."

Barker's methodology was used for rudimentary work by Coy (2001)[49] and Wright (2002)[50]. The Barker derivation stems from a less-intuitive (but more general) definition of the BGF than the memory function or impulse-response definition. Therefore the reasons behind it are inherently harder to understand physically and thus it is not used in this thesis.

[49] Coy (2001) performed a sensitivity analysis for a simple bimodal distribution of block sizes, for spherical geometry and also for spherical mixed with slab geometry.

[50] Wright (2002) performed a sensitivity analysis for different BGFs including ones with a discrete (bimodal) distribution. He observed an inexplicable crash during calibration (perhaps due to non-uniqueness invited by allowing the fraction of a particular size to float) and abandoned the method thereafter.

7.3.6 Choice of distribution of characteristic parameters

Different distributions have been investigated by various authors, as shown in Table 7.3. A good review of different distributions in relation to fractured rock appears in Bonnet et al. (2001).

Table 7.3: Choice of distributions of different parameters in the literature

Distribution	Author	Parameter
Uniform	Valocci (1990)	a
Lognormal	Haggerty and Gorelick, 1998;	$1/t_{cb}$
	Cunningham and Roberts, 1998;	t_{cb}
	Rameson, 1985; Haggerty et al., 2000	κ
Power-law	Haggerty et al., 2000;	$1/t_{cb}$
Gaussian	Rameson, 1985 ;	a
Gamma	Cunningham et al., 1997 ;	D_a
Discrete (bi-modal)	Haggerty and Gorelick, 1998 ;	$1/t_{cb}$
Multi-modal	Rameson, 1985 ;	κ (7-classes)
	Haggerty and Gorelick, 1995;	κ
	Cunningham and Roberts, 1998	D_a / a^2

Out of these distributions it is the **lognormal** which is most immediately conceptually appealing. Lognormal distributions tend to arise through the CLT being applied to multiplicative variability - a distinct possibility for diffusion time scales where D_a and a^2 vary multiplicatively. The entire pdf predicts positive values for the parameter in question.

Power-laws of timescales have been invoked in relation to fractal systems (Haggerty et al., 2000), which may have some resemblance to fractured rock (Berkowitz, 2002; Liu et al., 2004). They also arise in the context of continuous-time random walk models (Berkowitz and Scher, 1995) and in systems with diffusive or rate-limited sorption processes that are power-law distributed. Unless the distribution is truncated, there is no mean timescale.

The **gamma** distribution is less easy to relate to any underlying process and may simply be a reasonable empirical fit (it looks a little like the lognormal distribution, especially for large variance). Cunningham et al. (1997) showed how a gamma distribution can yield an analytical

solution to the governing equations, making it no more computationally demanding than for a single-diffusion model.

The **Gaussian** distribution is explicitly excluded by some authors as a non-starter, since without truncation it will predict negative transfer times, distances or exchange parameters. A Gaussian that was almost zero at $z = 0$ and then truncated might still be a reasonable model, but allowing the standard deviation to float, as a fitting parameter say, would risk calibration of what might be physically unlikely distributions.

The **uniform distribution** suffers from the limitation that, for a given mean, there is a maximum spread beyond which negative values would be included. The scope for influencing the BTC is lower than the potentially large influence a gamma or lognormal could be allowed to make.

Multi-modal distributions may be inherent to certain systems, but may also be the most appropriate model given data in the form of size classes (perhaps by sorting or sieving).

Continuous distributions, particularly for large spreads of scales, are more difficult to distinguish. Haggerty et al. (2000) noted that, despite a good lognormal fit to tracer data which spans over five orders of magnitude, on the basis on the BTC alone it is not possible to rule out other density functions including gamma or power-law densities.

These models clearly need strong constraint by data and logical application of possible parameter ranges. Even within the constraint, a fairly wide range of models may fit the data. For calibration against BTC datasets the problem of unique model identification arises again and predictive approaches that include a range of models may be necessary.

7.3.7 Least squares

It would be an interesting future exercise to fit a distributed-diffusion DP model with a lumped (i.e. single characteristic time) DP model. This could compare the effect of changing assumed distributions of distributed diffusion coefficients on the parameters of an unweighted least squares by a lumped DP model. This may indicate what happens if there is interpretation of a system that is really a distribution of scales by a single-scale model. Perhaps this may turn out to be a way to explain the 'scale effect' of diffusion coefficients, posited by Liu et al. (2004).

7.3.8 Moments

Cunningham and Roberts (1998) gave a table of moments for spherical and slab geometries. In the same manner as was applied to the findings of Sanchez-Vila and Carrera (2004), a more generalised moment treatment is given here.

Expressing $\mathrm{B}\left(\sqrt{st_{cb}}\right)$ as a McLaurin series in s gives:

$$\begin{aligned}\mathrm{B}\left(\sqrt{st_{cb}}\right) &= \mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} + s\frac{d}{ds}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} + \frac{s^2}{2!}\frac{d^2}{ds^2}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} \\ &+ \frac{s^3}{3!}\frac{d^3}{ds^3}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} + \frac{s^4}{4!}\frac{d^4}{ds^4}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} + \ldots\end{aligned} \tag{7.5}$$

Differentiating this series n times w.r.t. s yields series expressions for $\frac{d^n}{ds^n}\mathrm{B}\left(\sqrt{st_{cb}}\right)$. Since $\hat{\mathrm{B}}^{(n)}(0,t_{cb}) = \frac{d^n}{ds^n}\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} = t_{cb}^n\frac{d^n}{dS^n}\mathrm{B}\left(\sqrt{S}\right)\Big|_{S=0} = t_{cb}^{(n)}\hat{\mathrm{B}}^{(n)}(0)$ and $\mathrm{B}\left(\sqrt{st_{cb}}\right)\Big|_{s=0} = 1$, the integral for distributing diffusion (Equation (7.3)) can therefore be written as:

$$\int_0^\infty p(t_{cb})\left(1 + st_{cb}\hat{\mathrm{B}}'(0) + \frac{s^2t_{cb}^2}{2!}\hat{\mathrm{B}}''(0) + \frac{s^3t_{cb}^3}{3!}\hat{\mathrm{B}}'''(0) + \frac{s^4t_{cb}^4}{4!}\hat{\mathrm{B}}''''(0) + \ldots\right)dt_{cb} \tag{7.6}$$

Evaluating Equation (7.6) and defining μ_t and σ_t as the first and second moments of $p(t_{cb})$, $\int p(t_{cb})\,t_{cb}dt_{cb} = \mu_t$ and $\int p(t_{cb})(t_{cb} - \mu_t)^2\,dt_{cb} = \sigma_t^2$ gives:

$$1 + s\mu_t\hat{\mathrm{B}}'(0) + \frac{s^2}{2!}\left(\sigma_t^2 + \mu_t^2\right)\hat{\mathrm{B}}''(0) + \ldots \tag{7.7}$$

Inserting this series solution of the integral of times into Equation (7.4) and applying the method of moments yields the temporal moments given in Table 7.4, which are compared for the single block-time case which was discussed in Chapter 5.

Table 7.4: Moment comparison of DP with discrete t_{cb} versus DP with distribution of t_{cb}.

Discrete time (lumped) DP model	Distributed time (locally stochastic) DP model
$\mu_1' = t_a(1+\sigma)$	$\mu_1' = t_a(1+\sigma)$
$\mu_2 = 2t_at_d(1+\sigma)^2 - 2t_a\sigma t_{cb}\hat{\mathrm{B}}'(0)$	$\mu_2 = 2t_at_d(1+\sigma)^2 - 2t_a\sigma\mu_t\hat{\mathrm{B}}'(0)$
$\mu_3 = 12t_at_d^2(1+\sigma)^3 + \frac{4}{5}t_at_d\sigma(1+\sigma)t_{cb} + 3t_a\sigma t_{cb}^2\hat{\mathrm{B}}''(0)$	$\mu_3 = 12t_at_d^2(1+\sigma)^3 + \frac{4}{5}t_at_d\sigma(1+\sigma)\mu_t + 3t_a\sigma(\sigma_t^2+\mu_t^2)\hat{\mathrm{B}}''(0)$

From Table 7.4 it is clear that regardless of the distributions, the first moments always match. Setting $\mu_t = t_{cb}$ matches the second moments (Cunningham and Roberts, 1998). If $\mu_t = t_{cb}$,

then the third moment does not match (and t_{cb}^2 is replaced by $\sigma_t^2 + t_{cb}^2$). This translates qualitatively to saying that the tendency to 'tailing' is increased by the existence of a distribution of times. In practise this means that the time to flush the (potentially crucial) last 5% or so of solute can be dramatically influenced by the distribution of times. Cunningham and Roberts, 1998 illustrate how the slowest of a discrete distribution of times (corresponding to grain sizes) comes to dominate the 95% mass removal time for the Borden sands.

If however the integral was not written over t but over size, a, then matching would not occur when the mean, $\mathrm{E}(a)$, was set to a (Cunningham and Roberts, 1998).

Haggerty et al., 2000 showed for the MIM model BGF that the effective $\hat{\kappa}$ to match the first temporal moment is $\frac{1}{\hat{t}_M} = \hat{\kappa} = \frac{\sigma}{\int_0^\infty \frac{b(\kappa)}{\kappa} d\kappa}$, where $b(\kappa)$ is the pdf of MIM exchange parameters. So the effective MIM parameter is defined as a harmonic mean of the density function. Haggerty et al. (2000) called this the 'best' choice of a single parameter, in terms of the fact that it matches the zeroth, first and second moments of the BTC. Of course 'best' really depends on what is required.

7.3.9 Illustration of late-time BTCs

There follows a series of example simulations for different distributions, with the aim of demonstrating broad patterns. The codes are detailed in Appendix 1.

7.3.9.1 Bimodal

A flushing response to a bimodal distribution of block times compared to a single block time is given in Figure 7.7. The result is a characteristically 'benched' response. Note, however that whatever the different times are set to be it is not possible to generate multimodal arrivals in the manner of a dual permeability response. The impulse response is shown in Figure 7.8, demonstrating a substantial peak at a similar time to the 'bench' in the flushing response.

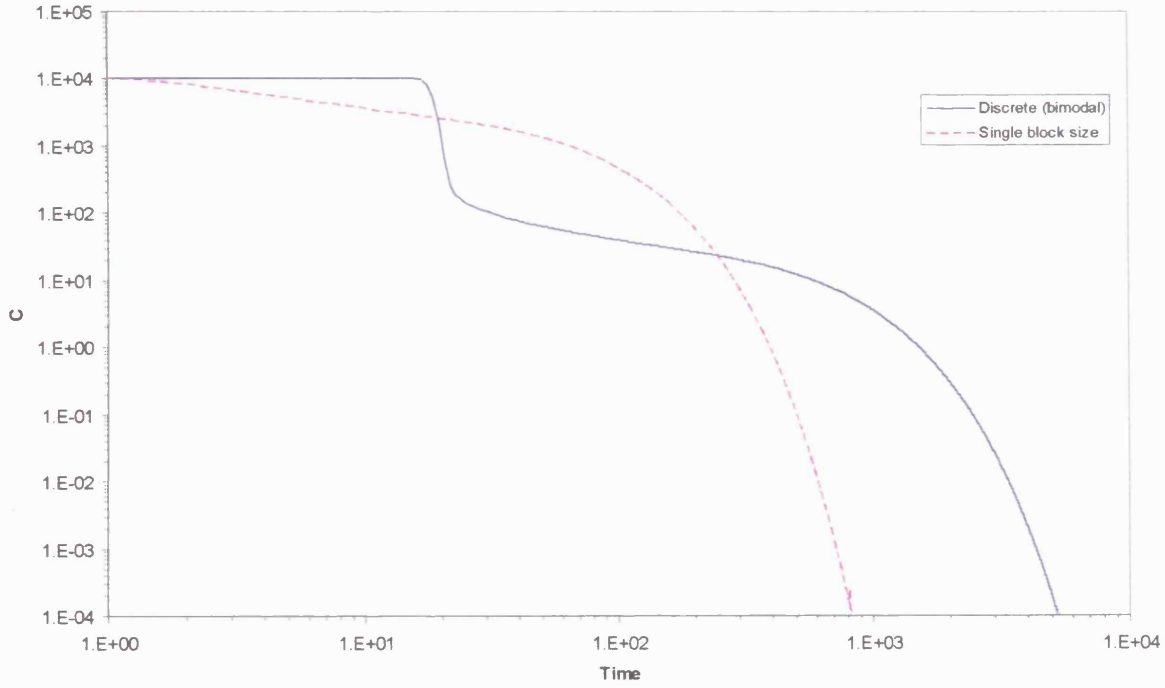


Figure 7.7: Bimodal block time distribution flushing response for DP(slab) geometry($t_{cb1} = 0.1$ with $P = 0.9$ and $t_{cb2} = 1000$ with $P = 0.1$ compared to single $t_{cb} = 100.09$. $V = 1$, $z = 1$, $\sigma = 20$, $\theta_m = 0.01$, $C_0 = 10000$).

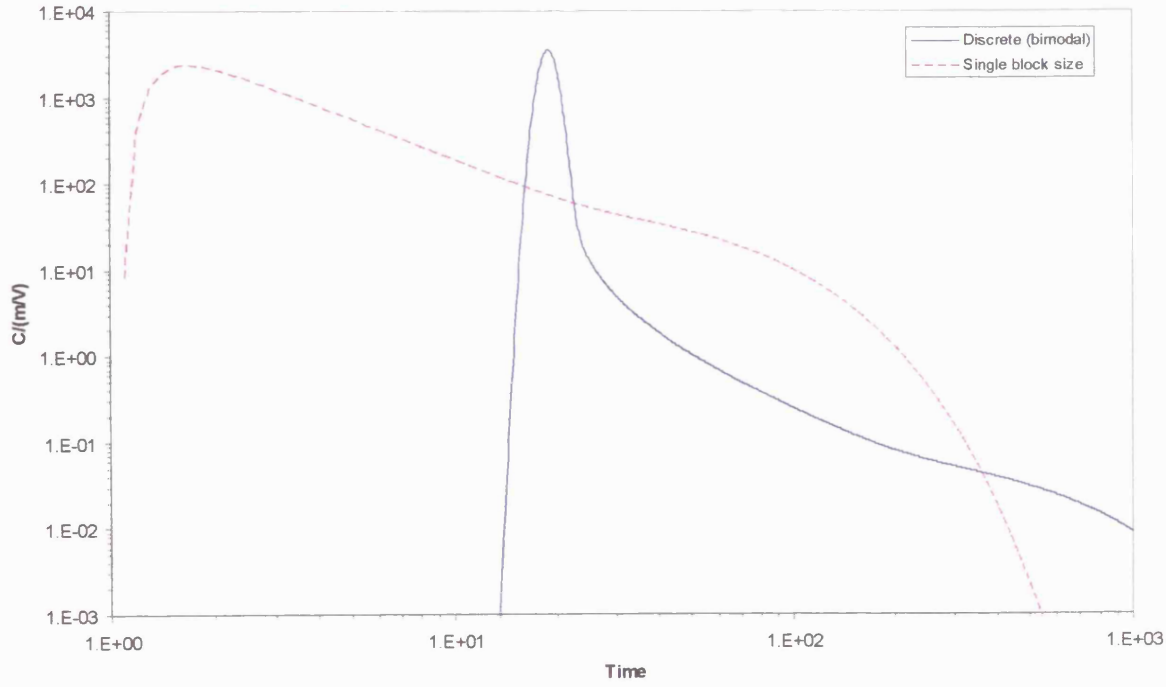


Figure 7.8: Bimodal block time distribution impulse response (same parameters as Figure 7.7).

7.3.9.2 Uniform

The flushing response to a uniform distribution of block sizes compared to a single block is given in Figure 7.9. The result is a relatively small perturbation to the late-time tail, without

dramatic change to the shape of the BTC. The width of the distribution is parameterised by the half-width, HW. There is minimal difference between the simulations down to $C/C_0 = 0.1$. Thereafter only the $HW = 150$ case diverges significantly, having a shallower tail.

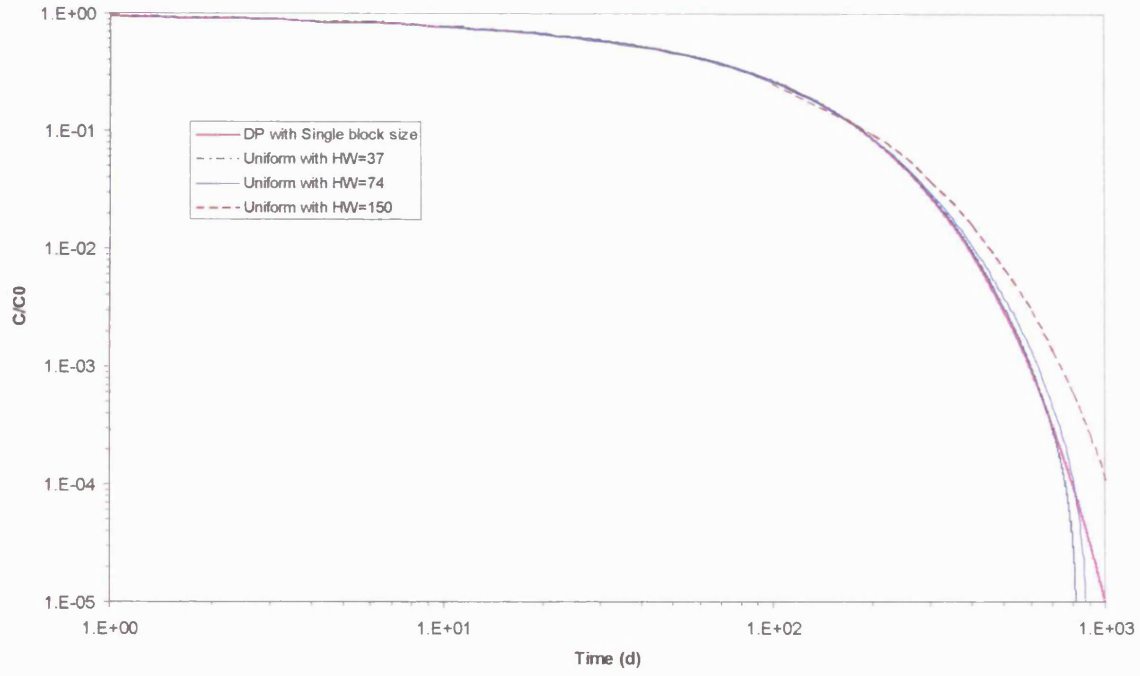


Figure 7.9: Flushing BTC for uniform block time distribution -central $t_{cb} = 200$ and the single DP simulation is for $t_{cb} = 200$ ($V = 1$, $L = 1$, $\sigma = 20$, $\theta_m = 0.01$, $C_0 = 1$ and different values of uniform distribution HW are compared).

Valocci (1990) assumed a uniform distribution of aggregate sizes and therefore inferred the MIM time distribution. It is worth remarking that such transformations do not necessarily keep the same average MIM time, so the moment-matching properties found by Cunningham and Roberts (1998) and illustrated in Section 7.3.8, no longer hold.

Valocci (1990) demonstrated a sensitivity whereby the first moment of the BGF for the distributed case was compared to the first moment of a single-site DP(spherical) model (thereby implicitly comparing the second centred moment of the BTC too). A small change (8%) in the ratio of moments was found even when the aggregate size distribution varied from 50% to 150% of the average size. Valocci (1990) also made a single visual comparison between the models which also proved to be similar. He concluded that the moments method therefore is a useful tool for assessing such sensitivities without having to run many repeated transport simulations and compare the fits. The more general treatment by Cunningham and Roberts (1998) now supersedes that work.

7.3.9.3 Lognormal

The flushing response to a lognormal distribution of block sizes compared to a single block size is given in Figure 7.10. The distribution has a profound impact on the shape of the BTC. The larger the spread the earlier the flushing effect and the shallower the mid-time log-log gradient (which is almost linear). This distribution is given by:

$$p(t_{cb}) = \frac{1}{t_{cb} SD\sqrt{2\pi}} \exp\left(\frac{-(\ln t_{cb} - \mu)^2}{2SD^2} \right) \tag{7.8}$$

where μ is the mean of $\ln t_{cb}$ and the lognormal standard deviation is SD. Note that the mean of t_{cb} is $\bar{t}_{cb} = \exp\left(\mu + \mathrm{SD}^2/2\right)$, hence $\mu = \ln\left(\bar{t}_{cb}\right) - \mathrm{SD}^2/2$.

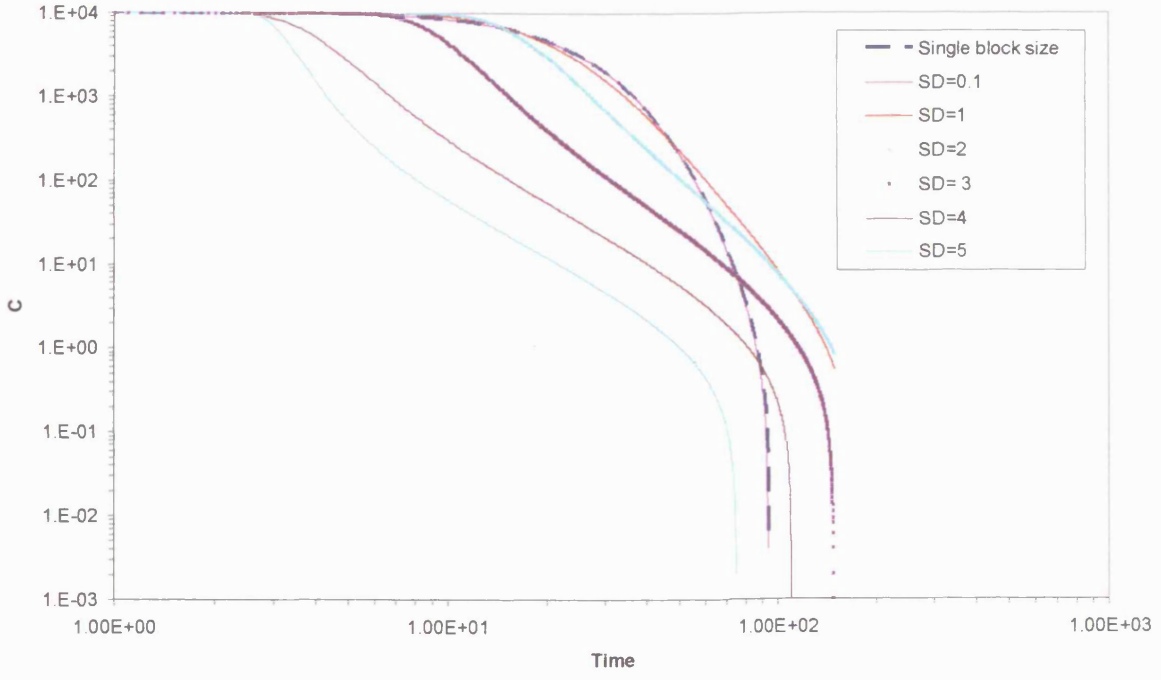


Figure 7.10: Flushing BTC for lognormal block time distribution (mean $\bar{t}_{cb} = 10$ and the single DP simulation is for $t_{cb} = 10$, $V = 1$, $L = 1$, $\sigma = 20$, $\theta_m = 0.01$ and $C_0 = 10000$ and lognormal standard deviation, SD is varied).

This graph exhibits an unexpected behaviour[51]. The concentration in the tail increases up to approximately $SD = 1$ and then begins to decline in a systematic way (the point at which $C < 10000$ then becomes ever earlier as SD increases).

[51] Further investigation of the discretisation and truncation confirmed that this is not due to numerical error.

7.3.9.4 Late-time summary (gradients)

Haggerty et al. (1999, 2000) has been able to achieve analytical solutions for late-time concentrations by making an approximation, giving:

$$C_f(t) \approx t_a \left(C_0 \sigma \mathrm{B}_t(t) - (m_0 / V_m) \sigma \frac{\partial \mathrm{B}_t(t)}{\partial t} \right) \tag{7.9}$$

The assumptions behind this are detailed in Chapter 5, where it is applied to single characteristic time systems. It applies equally to distributed systems, with the simple caveat that B is replaced by the distributed integral expression.

A distribution of diffusion times can eliminate the -3/2 log-log gradient for a single block size. This is because, as long as the distribution includes small enough times, there is never a point reached at which the entire system appears to be 'infinitely large'.

Some key results extracted from Haggerty et al. (2000), which was written with particular attention to MIM timescales are outlined below:

- A **Gamma** density function of MIM terms gives a late-time approximation of $C = (m_0 / V_m) t_a \sigma \gamma^2 \frac{\eta(\eta+1)}{(\gamma t+1)^{\eta+2}}$. For $t >> \frac{1}{\gamma}$ This has a log-log straight line slope, of $k = \eta + 2$. If k<3, then $\eta < 1$ gives infinite residence times.[52]
- A **Lognormal** density function of t_{cb} if σ>3 gives a slope between -2 and -3 (but not strictly a straight-line slope). For very large σ the lognormal distribution becomes very flat (approximately log-uniform) and a log-log gradient of -2 is approached.
- A **Power-law** density function of MIM terms with truncated power-distributions (i.e. with max and min limits) gives $C \sim t^{-k}$, thereby giving straight (log-log) plots, with slope k (where k is the power-law exponent). Note that the slope for an initially solute-filled system being flushed is given by $C \sim t^{-k+1}$.

In terms of the 'classical' infinite log-log linear gradient of -3/2 (indicative of an 'infinite' block system), the distributed approach need not completely preclude such a gradient. If there is a range in block sizes, with a distinct minimum size, then there may still be a log-log gradient of -3/2. The -3/2 gradient must be remembered as an indication of apparently infinite blocks, not as an indication of single block size.

[52] γ is the scale parameter and η is the shape parameter for a gamma distribution. Several studies have established these parameters by calibration to data (Connaughton et al., 1993; Pedit and Miller, 1994; Deitsch et al., 1998; Werth et al., 1997). $\eta > 1$ can occur but is relatively rare. This implies that most often the gradient is between -2 and -3.

For MIM with distributed block sizes log-lin linear for bi-modal will be $-\kappa$ (for κ corresponding to the longest MIM time t_M).

7.3.10 Locally stochastic dual porosity conclusions

The findings for locally stochastic DP representations using each of the methods are as follows:

Fitting

So long as the harmonic mean of the block times is correctly set, there is potentially very little observable change to concentrations shown on linear axes.

Moments

Moment analysis also shows no change for the first two temporal moments of the BTC, so long as the mean of the distribution is matched to the equivalent single characteristic time.

Late-time

Gradients of concentration against time can be substantially altered. Characteristically stepped gradients emerge for bimodal distributions, the -3/2 'infinite' slope can disappear, and the gradient can be substantially altered by different distributions. There is not a unique mapping from one distribution to a particular gradient.

7.4 *Locally-stochastic sorption representation*

7.4.1 Incorporating heterogeneity

Porous media can exhibit a wide range of different sorption surfaces, perhaps due to variability at the mineral scale or larger scale heterogeneities resulting from the presence of different soils, strata or mixtures of component materials (for example, in Municipal Solid Waste). Figure 7.11 illustrates this concept schematically.

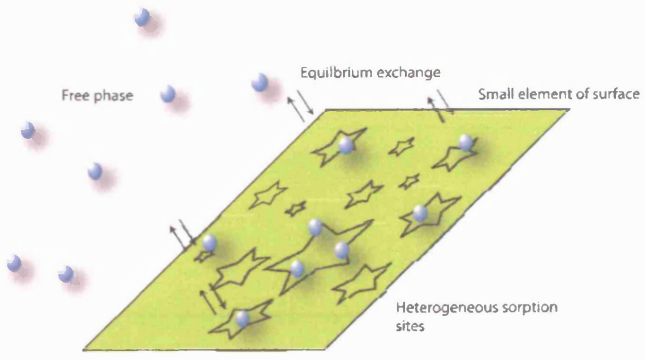


Figure 7.11: Sorption at a range of different sites.

An approach to sorption site heterogeneity is given in Kinniburgh et al. (1983), whereby a 'composite' isotherm is built up from a distribution of groups of sites. In general they defined a weighting function, which they call the 'Site Affinity Distribution Function' (SADF). This represents the fraction of sites with a particular site affinity. They compared fits to data with different underlying isotherms and allowed the SADF to vary in order to provide the best fit. The nomenclature is adapted to be consistent with this thesis. So:

$$\frac{C^*}{\beta C} = \int_0^\infty w(\alpha)\mathrm{f}(C,\alpha)d\alpha \tag{7.10}$$

A discrete version of Langmuir would then be $\frac{C^*}{C} = \sum_{i=1}^{k} \frac{\alpha_i \beta_i}{1+C\alpha_i}$, where $\beta_i = w(\alpha_i)\beta_{tot}$ and $\beta_{tot} = \int_0^\infty w(\beta)\beta d\beta$ is the total adsorption capacity.

A two site Langmuir model ($i = 2$) would therefore be $\frac{C^*}{C} = \frac{\alpha_1\beta_1}{1+\alpha_1 C} + \frac{\alpha_2\beta_2}{1+\alpha_2 C}$. This gives retardation, $R = 1 + \frac{B_d}{\theta}\left(\frac{\alpha_1\beta_1}{(1+\alpha_1 C)^2} + \frac{\alpha_2\beta_2}{(1+\alpha_2 C)^2}\right)$.

An expression for general retardation can be derived (Hinz et al., 1994), using $R = 1 + \frac{B_d}{\theta}\frac{\partial C^*}{\partial C}$ to give:

$$R = 1 + \frac{B_d}{\theta}\frac{\partial}{\partial C}\left(\beta C \int_0^\infty w(\alpha) f(C,\alpha)d\alpha\right) \tag{7.11}$$

For the Langmuir isotherm, this yields:

$$R = 1 + \frac{B_d \beta}{\theta}\int_0^\infty w(\alpha)\frac{\alpha}{(1+\alpha C)^2}d\alpha \tag{7.12}$$

Expressing the heterogeneity in this way is equivalent to the form of parallel summation given by Haggerty and Gorelick (1995), which was applied in the case of a MIM exchange process. It is appealing here, as it gives a common thread to the heterogeneity representation for AD, DP and sorption. In common with the distributed diffusion approaches (Chapter 7) is the difficulty of constraining the distribution of stores, yet there is a potentially large role played by this distribution in determining the tail of a flushing BTC.

In terms of a practical application of these integrals, there may be a case for producing composite isotherms based on known component mixtures, such as well-characterised wastes. This would involve distributing sites over a much bigger scale than the molecular-scale over

which distributions were taken by Kinniburgh et al. (1983), who worked with a more formal statistical thermodynamics approach.

In terms of the predicted statistics of the BTCs with such isotherms, Bosma et al. (1996) observed that,

> "no analytical expressions are yet available for the development of the various moments as a function of time".

However, Dimitrova et al. (2002) examined temporal moments by numerical methods. No analytical expressions for late-time have been obtained, nor has any literature comparing the fits of different heterogeneous representations been found.

7.4.2 Further heterogeneous sorption Literature

Hinz et al. (1994) noticed considerable differences in BTCs when applying different isotherms fitted to the same data and provided some discussion of the effect of different assumptions. They showed that the higher the average of the SADF the more extensive the tailing. This is analogous to the distributions of t_{cb} with more contribution from the larger time fraction exhibiting greater tailing.

A very clear and practical contribution is from Kinniburgh et al. (1983), who suggested that if the simplest isotherms do not model the data correctly, then more complicated expressions may be more appropriate. They suggest that Langmuir and Freundlich isotherms have enjoyed rather more 'success' than they deserve, since close examination of the data often reveals systematic deviations from the isotherms. They demonstrated specific distributions of activation energies (SADFs) that, when integrated over a basic Langmuir, could produce specific isotherms.

Kinniburgh et al. (1983) described the Langmuir-Freundlich isotherm as $f(C)=\frac{\beta(\alpha C)^m}{1+(\alpha C)^m}$,

where m is a 'heterogeneity constant' that characterises the range in adsorption energies. For $m=1$ (for an infinitely narrow distribution) this collapses to a Langmuir. For $\alpha C \ll 1$, this collapses to a Freundlich isotherm. This model seems to suggest that the Freundlich model might be valid for very low C, but this is not what is often measured.

Kinniburgh (1986) further discussed some 'general purpose' isotherms, which are described as largely empirical models. This added the Dubinin-Radushkevich isotherm to the multi-site and modified Langmuir and Tóth isotherms, which had all appeared in the earlier paper. How and

whether these models collapse to the simpler Langmuir and Freundlich isotherms is of interest. These 'general purpose' isotherms are said to all be "well behaved", in as much as they become linear at small C and approach a maximum concentration.

There must be a physical limit to the sorption capacity. Only the Langmuir model out of the four mentioned here inherently limits the maximum capacity. The other isotherms must therefore have upper bounds set for them.

7.5 *Summary*

A very simple common thread is identified that appears under different names and is widely spread through the transport literature. The concept is to establish (by measurement, calibration or assumption) distributions of key 'unit' processes. This is conceptually simpler than attempting to represent the full spatial variability, which is also often very difficult to characterise. This represents a compromise or middle level of heterogeneity representation, perhaps for situations when simple process models are inadequate or non-conservative, but the full 3D reality is too complicated to model.

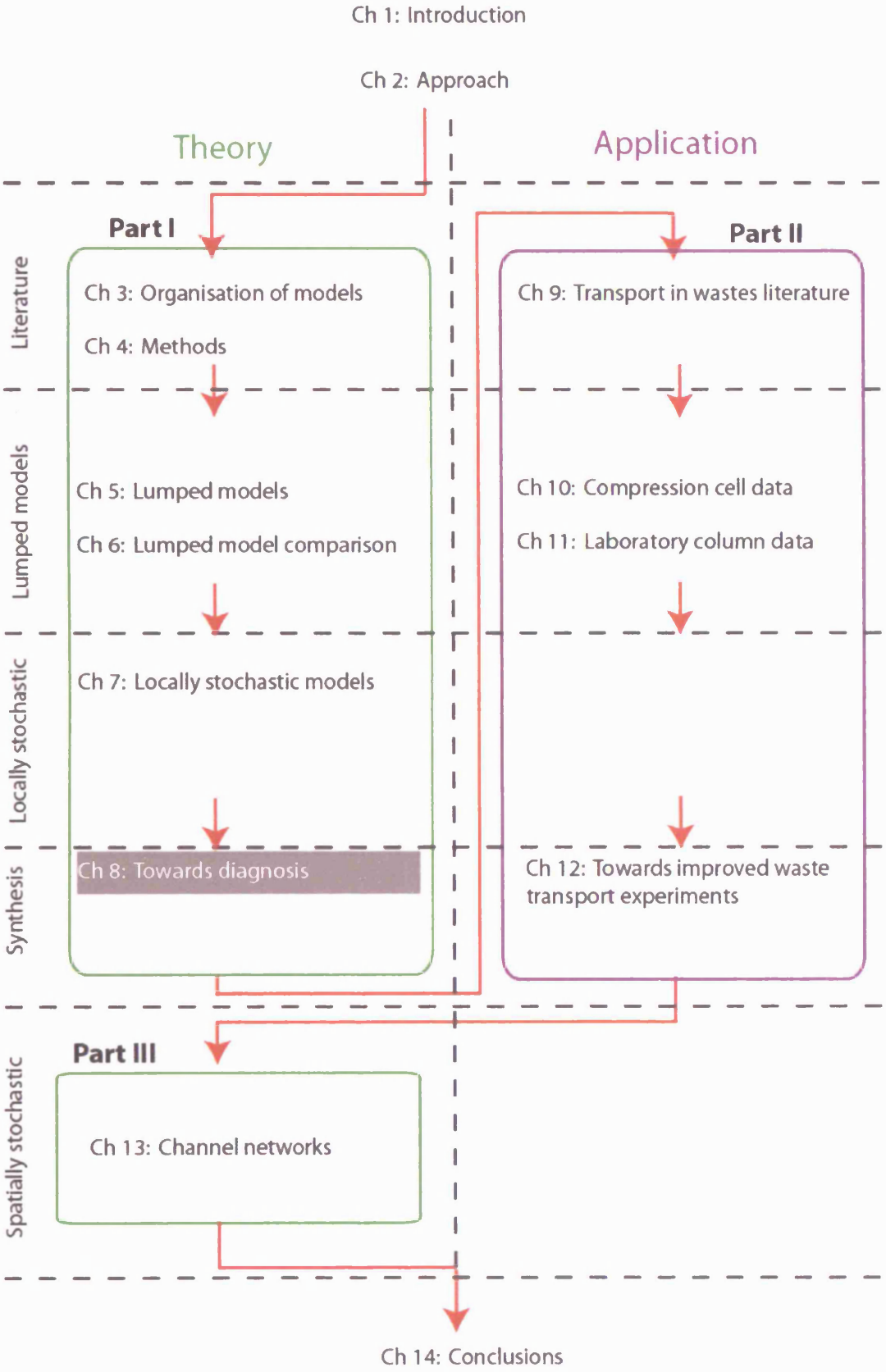
The selected representations are as follows:

AD is given by $\hat{C} = \int_0^\infty C(t_A)p(t_A)\,dt_A$;

DP is given by $\mathrm{B}(t) = \int_0^\infty p(t_{cb})\mathrm{B}\left(t,t_{cb}\right)dt_{cb}$;

And sorption is given by $\frac{C^*}{\beta C} = \int_0^\infty w(\alpha)\mathrm{f}\left(C,\alpha\right)d\alpha$

There is insufficient space to present a comprehensive comparison of each model, as examined by a range of measures in the manner of Chapter 6. This therefore completes the work towards Objective 1, leaving space for future comparative work. In the next chapter the insights gained from examining locally stochastic models are used in conjunction with the results for lumped models in order to enhance process diagnosis and find improved experimental methods.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Lumped models
Locally stochastic
Synthesis
Spatially stochastic
Ch 3: Organisation of models
Ch 4: Methods
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 7: Locally stochastic models
Ch 8: Towards diagnosis
Ch 9: Transport in wastes literature
Ch 10: Compression cell data
Ch 11: Laboratory column data
Ch 12: Towards improved waste transport experiments
Part III
Ch 13: Channel networks
Ch 14: Conclusions

8 Towards diagnosis

The purpose of Part I of this thesis has been to analyse the behaviours of different models in order to identify diagnostic traits. Three different BTC analysis methods were selected in Chapter 4 (moments, late-time gradients and fitting). Lumped models were then analysed in their own right using these methods in Chapter 5 and then compared in Chapter 6. A wider 'family' of more heterogeneous locally stochastic models was then examined in Chapter 7. It is important to 'drill down' to the detailed level of Chapters 5-7, however it is also necessary to examine diagnosis in a broader way. The latter is the aim of this current chapter.

There are different aspects to consider:

- What do different measures reveal about each process or process-combination?
- In what ways might different processes be confused? What apparent parameter scaling relationships exist if the system really resembles one process but is assumed to be another?
- Do the different measures used here give the same picture?

This chapter is structured in terms of diagnosis based on each of the three measures of concentration data used in the preceding chapters (see Figure 8.1). The relatively uninformative peak time is then briefly considered (in part to explain why it has not been selected as one of the measures). Next, enhanced experimental methods are considered in the light of possible confusion between processes. Two specific methods to increase diagnostic ability are considered: intermittent flow and multiple tracers. Finally, the diagnostic advantages of using a top-hat solute input function are presented.

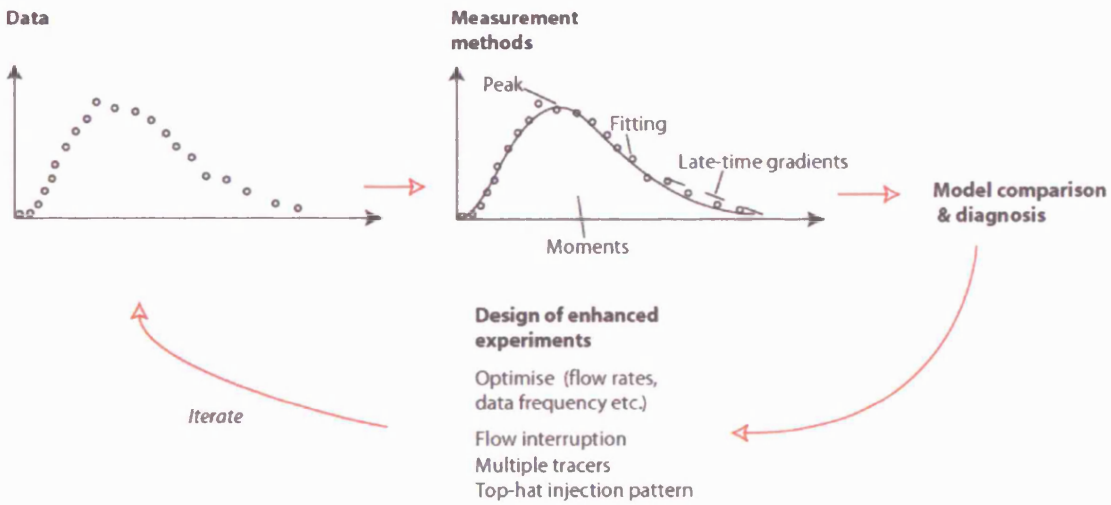


Figure 8.1: Schematic of diagnostic approach in Chapter 8.

Because the 'pool' of possible models is virtually infinite, a comprehensive table of all processes would be unfeasible. Processes can only be eliminated but not positively identified (in keeping with a Popperian rather than Positivist approach). Therefore, strictly, if the term 'diagnosis' is used to mean correct identification of the underlying processes, then diagnosis is impossible. Instead the term here refers to a *journey* towards better understanding; one which increasingly improves an approximate understanding. Ultimately the diagnostic traits of a much wider range of models need to be categorised and it is intended that the work here would be supplemented in the future. The work presented here provides a basis for which a future process-diagnosis expert system could be based. It is this progression towards a diagnostic methodology that forms the contribution of this chapter.

8.1 Late time gradients

8.1.1 Introduction

The effect of certain key factors must be considered, particularly:

- Location of boundaries.
- The influence of common additional transfer functions in series with the main process model (especially the mixing transfer function).
- Flushing and impulse responses.

Chapter 5 predominantly dealt with semi-infinite temporal impulse responses, which have particular relevance for modelling column experiments. However, the results in the instance of a spatial impulse between infinite up- and down-stream boundaries (for example a point spillage in an aquifer) are of interest for a wider range of problems than the focus here. This second condition is touched on here in order to highlight the fact that the assumed boundary conditions can have profound effects on prediction. The influence of an inline mixer is also incorporated into the following analysis, because it is commonly encountered. Also included is the effect of the initial and boundary conditions. The two primary sets of conditions that are analysed in this thesis are temporal impulses and the flushing of an initially uniform solute distribution with a step function of solute-free liquid.

8.1.2 Log-log constant gradients

Table 8.1 summarises the late-time gradient results from this thesis (predominantly from Chapter 5, but further results are also brought in from Appendix 2).

Whether there is an impulse input or a flushing condition makes no difference to the late time log-log slope for the AD or DP (with a single block time) models. However for the distributed-diffusion case, the boundary conditions can in some circumstances affect the gradient (e.g. the

slope for an initially solute-filled system being flushed is given by $C \sim t^{-k+1}$, whereas the slope for an impulse response is $C \sim t^{-k}$). For the log-lin slope the upstream boundary condition has a strong influence on the lumped AD model. For an initial spatial impulse with infinite up- and down-stream boundaries the log-log gradient is -1/2, whereas for a temporal impulse with a boundary at the origin and an infinite downstream boundary, the gradient is -3/2.

The ranges over which slope linearity occurs are given in Chapter 5. The ranges themselves may contribute to diagnosis where more than one process shares the observed gradient. The predictions of vertical offsets of the late-time produced by particular process changes are also a potential further (more subtle) tool that should be considered in the future.

8.1.3 Log-lin constant gradients

Table 8.1 shows particular examples found in this thesis for which processes give log-lin linear late-time gradients (under certain boundary and initial conditions).

Table 8.1: Whether processes have a log-lin linear gradient.

	Process	Comment	Reference
LIN LOG	AD impulse (temporal or spatial)	Linear sorption reduces the gradient to $-V/4\alpha R$	Chapter 5
	AD (flushing)	Gradient tends to $-V/4\alpha$ at late time	Chapter 5
	AD with an inline mixing cell	Gradient tends to $-V/4\alpha$ at late time (as long as mixing cell not too big)	Appendix 2
	DP (impulse and flushing)	Linear sorption reduces the gradient, for example to $-\pi^2/4t_{cb}R$ for a slab	Chapter 5
	DP(MIM) (impulse and flushing)	Linear sorption reduces the gradient to $-\theta_{im}/\kappa R$ for a slab	Chapter 5
	AD-DP model	This can have either a log-lin gradient determined by the DP process, but if exceptionally large mechanical dispersion is assumed the gradient can be the same as for an AD model	Chapter 5
NO LOG-LIN region	Freundlich sorption	Freundlich sorption as a 'lone process' only predicts log-log linearity. In conjunction with the AD model Freundlich sorption causes non-linear log-lin gradients	Chapter 5

	Process	Comment	Reference
	Doublet (without mechanical dispersion)	This advection-only model predicts only a log-log linearity. It would be interesting further work to examine the gradients for a doublet with AD and/or DP processes occurring in each stream-tube	Appendix 1
	DP(infinite)	Only a log-log linear region is predicted. Log-lin behaviour for a DP model begins for $t > t_{cb}$	Chapter 5
	Lognormal stochastic-advective model $C(t) = \frac{1}{\sqrt{2\pi\sigma t}} e^{\frac{-(\ln(tl/z)-\mu)^2}{2\sigma^2}}$	The gradient is $\frac{\mu'-\sigma^2}{t\sigma^2} - \frac{\ln(t)}{t\sigma^2}$, which is not linear. However, as long as $\frac{\mu'-\sigma^2}{t\sigma^2} \gg \frac{\ln(t)}{t\sigma^2}$ a linear log-log gradient would be given by $\frac{\mu'-\sigma^2}{\sigma^2}$	Chapter 8
	DP with multiple diffusion timescales	The existence of a period of log-log linearity in the presence of a distribution of timescales of exchange are special cases rather than the rule, so an absence of a linear gradient is entirely possible	Chapter 7

For the AD-DP model with relatively small mechanical dispersion the only impact on the BTC is at early-time and the late-time gradient is governed by the DP model.

Figure 8.2 compares the log-log linear gradients that are achieved by different models.

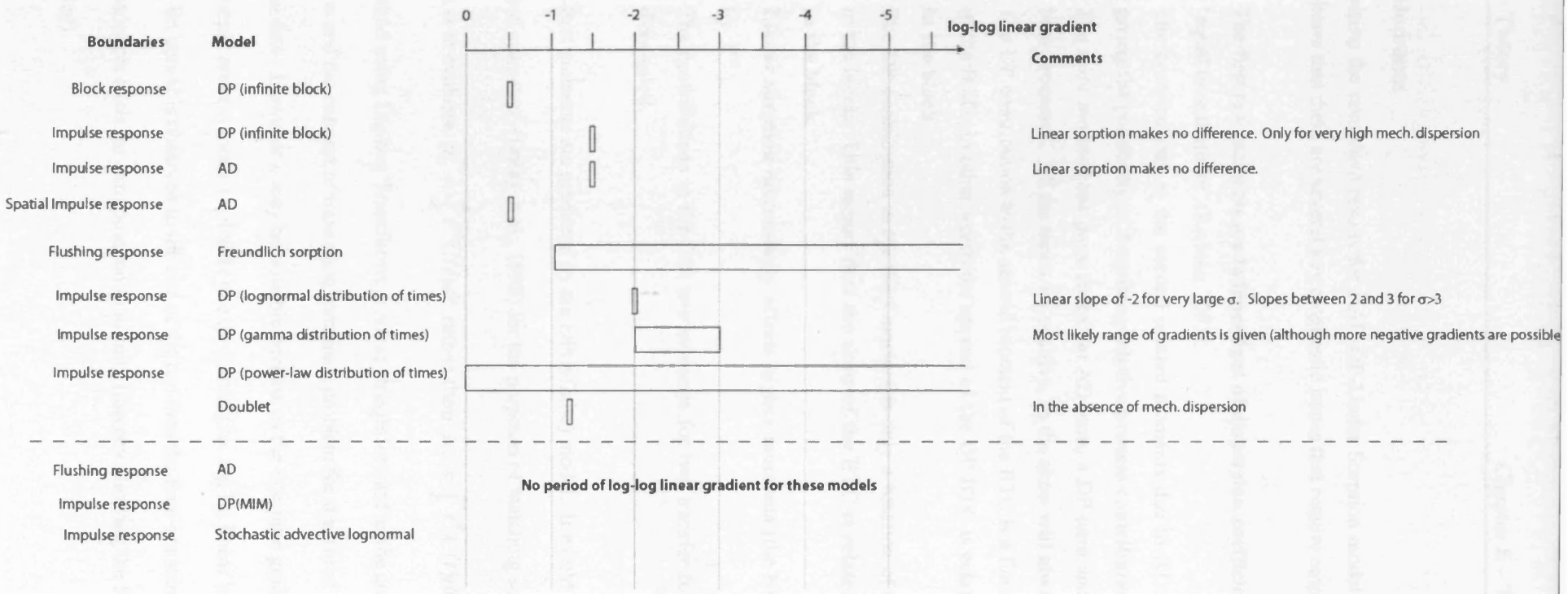


Figure 8.2: Log-log linear gradients for different process models that are discussed in this thesis.

8.2 Moments

Synthesising the combined results for a AD-DP-Linear Sorption model (see Sections 5.5.2 and 5.6.1) shows that there are several key diagnostic issues that require emphasis:

- The first raw moments are **independent of dispersion coefficient or BGF.** This is the 'equal time theorem' (Redner, 2001).
- The contributions to the second centred moments due to AD and DP are decoupled, giving the possibility of describing **relative process contributions** to spreading.
- The third moment has three terms: an AD term, a DP term and a cross-term involving both processes. All the terms are positive, so the **skew will always be positive.**
- The DP contribution to the second moment of the BTC is a function of the first moment of the BGF. In other words the **spread of the DP BTC is related to the average time in the block.**
- The DP contribution to the third moment is only a function of the second raw moment in the block. **This means that the skew of the BTC is related to the spread of time in the block.**
- **Linear sorption increasingly affects higher moments** (the Nth moment is multiplied by R^N).
- The contributions to the first two moments for two transfer functions in series can be **decoupled.**

Infinite first moments are produced by the DP(infinite) model. It would be possible to define 'truncated moments' (Delay et al., 1998) for the purposes of matching such systems. The basic concept is to evaluate $\mu_N^T = \int_0^T t^N C(t)dt$ rather than $\mu_N = \int_0^\infty t^N C(t)dt$. Implementation of this method using Laplace Transforms is not as straightforward as the conventional method of moments and the concept of truncation remains a problem for the use of the method of moments with real data. However it may be possible because as the late-time gradients work has shown, many simple process models collapse to exponential (i.e. log-lin linear behaviour) at late-time. If a log-lin period is observed to the end of the test then the time-constant of the gradient might be a reasonable basis for extrapolation to infinity (thereby allowing the full moment to be calculated).

Comparison of processes using moments

The detailed analysis in Chapter 6 was designed to reveal where processes might be confused. A few key diagnostic points must be emphasised here.

DP vs. AD: Exact moment matching of BTCs is possible for a restricted set of conditions between the AD and DP for the first three moments. The fourth moment is accurate for $Pe \gg 8$. The MIM model gives a worse match than the slab, cylinder and sphere models.

AD-DP vs. AD-MIM: Matching of the first BGF moment and the first two BTC moments is possible. However the second BGF and third BTC moments cannot be matched. This is due to the AD-DP cross-term in the third moment.

MIM vs. DP: Matching the first moment of the BGF is the same as the condition derived by Barker (1985b) who used a McLaurin expansion. Under this condition, the first and second moments of the BTC will also be matched.

8.3 Fitting

This topic has received much attention in the literature, but usually in the context of specific model pair comparisons. Few studies explore the use of synthetic noise or MC to examine the distribution of possible parameters and few studies examine the goodness-of-fit as well as presenting the parameters achieved at the best fit. Useful literature contributions have been organised in Table 6.1.

Comparing MIM and AD shows that for a match κ is linearly proportional to V/α.

8.3.1 Comparison of different models

DP (or MIM) fitted by AD.

A good fit was demonstrated with the entire BTC inside the 95% CI in an example where the first three moments were matched (Section 6.3.3). The CR was elliptical.

A scaling of $\frac{\alpha}{V} \propto t_{cb}$ for AD vs DP and $\frac{\alpha}{V} \propto \frac{1}{\kappa}$ for AD vs. MIM is observed by Bajracharya and Barry (1997) for systematic fitting over a range of parameters. This is the same as that found by matching the second moments of the same models and their log-lin gradients, although there is a difference in the constant of proportionality in each of the cases.

AD-Freundlich Sorption fitted by AD

A simple fitting exercise in Section 6.4.3 showed that the AD, MIM and DP models could be fitted to a step response for an AD-Freundlich model. This was not the case for a finite top-hat pulse, where the non-linear sorption created an asymmetry in the BTC which could not be matched, unless the pulse time was adjusted. This asymmetry of rising and falling limbs, above and beyond the asymmetry that is inherent for an AD BTC, can be turned into a diagnostic test.

AD-DP fitted by AD-MIM

This comparison has been comprehensively covered in the literature and is shown in Section 6.5 to be a function of the 'contact' time of the tracer (i.e. of the experimental conditions).

DP fitted by MIM

A simple example was used in Section 6.6.3 to demonstrate with choice of κ' to match the first moments of the BGF gives a better fit than the choice of κ' to match asymptotically. The choice of κ' is therefore a *pragmatic* one, depending on the problem in hand.

Distributed-time DP fitted by single-time DP

Haggerty et al. (2000) showed that if a system that has a particular power-law distribution of exchange timescales (that gives $C \sim t^{-k}$ and where $k \leq 3$) is interpreted using a single-time DP model, then the single characteristic time that is established will depend upon the timescale of the experiment.

8.4 Peak time (mode)

Although not a method selected for systematic analysis in this thesis it is still useful to include the peak time as a contributory diagnostic measure. The time and concentration of the peak may be of great practical interest (the 'worst' point of contamination if pollutant transport is being considered). However there are several problems with using the time or concentration of the peak to constrain the model.

Firstly, the peak is potentially the noisiest part of the BTC (say, for example when the error is proportional to concentration). Secondly, it is impossible to obtain the time of peak in Laplace space. It can be evaluated by differentiation for the special cases which are analytically invertible (e.g. the infinite block and the various AD solutions). For all other cases, the peak has to be found numerically (either by looking for a minimum, or by differentiating in Laplace space by multiplication by s and then root-finding).

The peak time for the lumped AD model with infinite boundaries and an initial solute distribution of $(m_0/V)\delta(z)$, described by $C = m_0 \frac{z}{\sqrt{4\pi Dt}} \exp\left[-\frac{(z-Vt)^2}{4Dt}\right]$, is given by:

$$t_{AD}|_{\max} = \sqrt{t_A^2 + t_D^2} - t_D \tag{8.1}$$

So, for $t_D \gg t_A$ $t_{AD}|_{\max} = 0$ and for $t_D = 0$, $t_{AD}|_{\max} \approx t_A$.

The peak time for AD with an infinite downstream boundary and an impulse input of solute, described by $C = (m_0/V)\frac{z}{\sqrt{4\pi Dt^3}} \exp\left[-\frac{(z-Vt)^2}{4Dt}\right]$ (Equation 5.7) is given by:

$$t_{AD}|_{\max} = \sqrt{t_A^2 + 9t_D^2} - 3t_D \tag{8.2}$$

For $t_D \gg t_A$ $t_{AD}|_{\max} = 0$. For $t_D = 0$, $t_{AD}|_{\max} \approx t_A$. In other words, the mean time for an AD system (which will be shown to equal t_A), gives an upper bound for the peak (conversely the peak is a lower bound for the mean) and the lower bound is the start time.

The peak time for DP with an infinite-block (see Section 5.3) is given by:

$$t_{DP(\text{inf})}|_{\max} = t_a\left[1 + \frac{t_a}{6t_{cf}}\right] \tag{8.3}$$

So, for $t_{cf} \gg t_a$ $t_{\text{DP(inf)}}|_{\max} \approx t_a$ and for $t_{cf} \ll t_a$, $t_{DP(\text{inf})}|_{\max} \to \infty$

For DP peak time, no analytical results are possible in the time domain, but numerically inverting a Laplace Transform solution gave $t_a \le t_{\text{DP}}|_{\max} \le t_A$ (note the upper limit in contrast to the infinite block upper limit). For AD-DP, also no analytical results are available in the time domain, but numerically inverting a Laplace Transform solution gave $0 \le t_{\text{ADDP}}|_{\max} \le t_A$.

It is highly tempting to use single measures to achieve some sort of parameter bounding. Caution must be taken in using the peak to define 'upper' or 'lower' bounds for processes or parameters, because the system may simply not behave as expected through intuition. For example it might be imagined that fitting the peak assuming no AD would give an upper bound for the contribution to the peak due to DP. Adding AD would be expected to lower the peak, but it can, under the right conditions, actually raise it (see Figure 5.28).

The peak time therefore provides a less diagnostic measure than might be assumed. It is arguably an over-used measure.

8.5 Experimental enhancement

In order to enhance the diagnostic potential of transport experiments there are a few methods that are worth consideration. In particular the use of intermittent flow (or 'flow interruption'), simultaneous use of multiple tracers and use of top-hat injection patterns instead of step inputs.

8.5.1 Intermittent flow

In systems such as column experiments there is potentially tight control of the flow rate. This presents an opportunity to test models under different flow rates (as discussed by Jury and Roth (1990)).

The limiting case for varying the flow rate is to test for what happens when there is a period of no flow at all. In this instance the mechanical dispersion process is presumed to completely stop operating, restarting with the same concentration distribution as before the stop. In contrast, the DP or MIM exchange models predict continued exchange between zones until concentration has equilibrated. Other non-equilibrium processes might continue, such as non-equilibrium sorption or decay, causing a change in observed concentration after the stop. This will either be more, if the immobile concentration was greater than the mobile, or less if the reverse is true.

Figure 8.2 shows the modelled concentrations for a MIM model. Note that the mobile concentration is greater than the immobile concentration in the rising limb of the BTC and remains so until a 'cross-over' following the peak in mobile concentration. Thereafter the immobile concentration exceeds the mobile concentration. The mobile and immobile concentrations are shown following a stop in flow that occurs after the 'cross-over' (i.e. when exchange from immobile to mobile zones occurs).

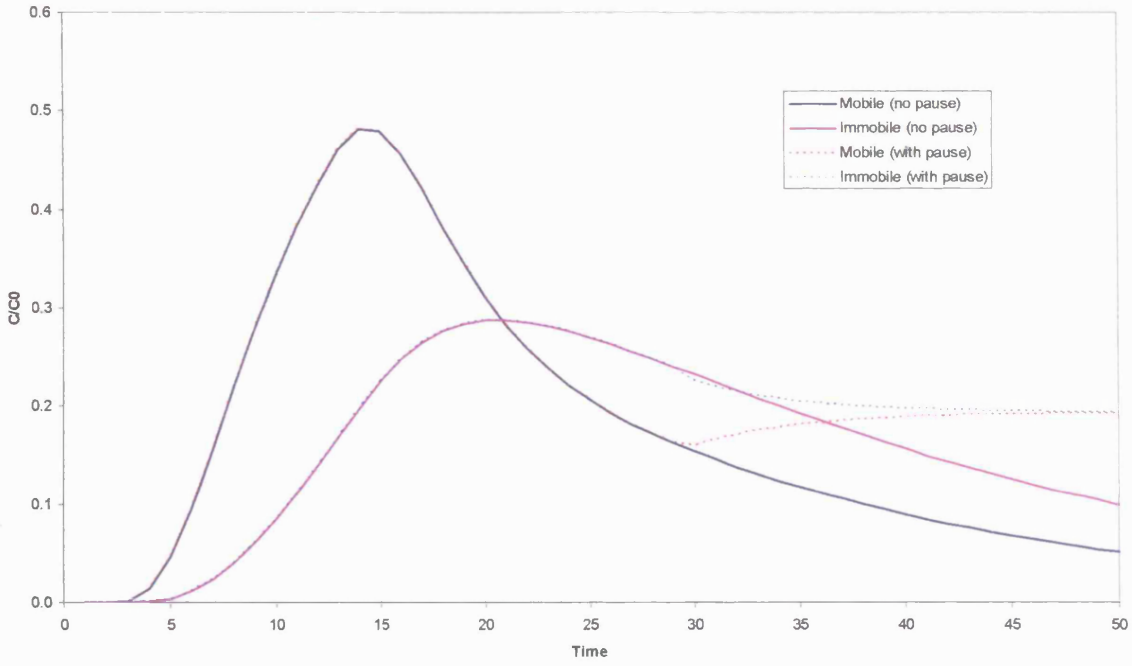


Figure 8.3: MIM mobile and immobile concentrations ($T_p = 10$, $\sigma = 1$, $V_m = 0.1$, $\kappa = 0.001$, $L = 1$). The dotted lines give concentrations that would occur if the flow were stopped at $T_s = 30$ and continuous lines give concentrations with no stop.

Model identification

Brusseau et al. (1997) discussed the explicit use of flow interruption to attempt to distinguish between particular process-pairs, namely: diffusive non-equilibrium exchange and heterogeneity (i.e. DP and AD); rate-limited (first order) sorption and non-linear sorption; transformation (i.e. decay) and sorption. Fortin (1997) used flow interruption to distinguish between rate-limited sorption and DP exchange.

Each of the model formulations, DP, MIM (whether representing sorption or diffusion) or decay will predict a drop in concentration if the test was stopped before the 'cross-over'. However, after the 'cross-over' only the DP and MIM models predict a rise in concentration following a stop, whereas decay predicts a drop in mobile concentration. A stop after the 'cross-over' enables discrimination between decay and DP or MIM effects. If there is a combination of DP and decay processes, the two effects can potentially counteract and confuse the analysis. Brusseau et al. (1997) compares the BTCs of different models.

Parameter estimation

Reedy et al. (1996) modelled a series of 1D transport tracer tests in a fractured saprolite with a MIM model. They demonstrated that a reasonable visual fit for the entire curve (except the stopped section) could be achieved for larger values of κ than the values achieved for the best-

fit including the stopped section. They therefore concluded that this gives additional constraint to parameter fitting.

Wright (2002) examined intermittent flow in the context of aquifer storage and recovery (ASR) using the DP-PULSE model. A typical pumping pattern was an injection of clean water into an initially uniform distribution of solute in both mobile and immobile phases, followed by a stop and then recovery by reversing the pumping direction and abstracting from the same well. This pattern was repeated several times. Wright (2002) was interested in optimising the configuration of a tracer test with injection, stop and withdrawal (an 'ECHO' test) for optimum parameter recovery in order to condition the operation of an ASR system. Operational research methodologies were employed (i.e. including cost and time as important considerations as well as the consideration of data-worth). As long as the pumping rate was the same as the abstraction rate, the model was invariant to both flow rate and geometry. This seems surprising from an intuitive point of view, though was predicted by Mackenzie and Barker (1998). Calibration against synthetic and real data showed that the first cycle gave t_{cf}, but revealed little about t_{cb} or σ. Increasing the cycle duration or running more cycles began to reveal t_{cb} with increasing accuracy.

In a theoretical study of column transport experiments, Altmann-Dieses et al. (2002) performed a general optimisation of the input function (piecewise constant q and C) and the sampling frequency with the aim of reducing the CIs for all calibrated parameters (based on unweighted least-squares). They applied this to an unsaturated soil model and used a non-linear optimisation routine, which massively improved the CIs on each parameter compared to simple constant input functions. The optimized input function was complicated, but the key feature was that both q and C went from maximum to zero.[53] It is not clear whether the solution was in a unique minimum, but the very low noise level that they used to generate synthetic data ($\sigma_N = 0.00005C_0$) will have ensured that an elliptical error surface existed.

Future work

When using flow-interruption in column experiments in order to enhance parameter recovery it would be useful to optimise the timing and duration of the stop. This could be done by using prior estimates of parameters (for example of t_{cb} and σ in a DP model) to generate a model prediction and then running MC simulations of the same model with random noise added to the first prediction. The CIs must be used cautiously because the error space may not be

[53] This is for an unsaturated model. It is not yet clear what input function is optimal for a saturated system.

ellipsoidal. Use of percentiles on a cdf would be one way of measuring the spread of possible parameter values that is not susceptible to this problem.

This method could be extended to the analysis of more general models (for example the AD-DP model). It would be very useful in this instance to optimise the stop in order to minimise the CIs for all the parameters. Although this is strictly a parameter estimation problem, the aim is also effective process identification, because it is often difficult to distinguish the contributions of AD and DP processes from a simple constant flow-rate BTC.

8.5.2 Multiple tracers

Inferring processes and parameters from a single tracer BTC can be affected by uniqueness questions. Repeated measurement under different conditions (say for changed flow rate or measurement at different distances) is a way to address this. However such approaches take longer, may be difficult to arrange in practical terms and potentially cost significantly more.

Simultaneous use of multiple tracers potentially overcomes non-uniqueness between models without requiring the experiment to be repeated. The method relies on the different tracers having well-described transport properties. The tracers should not participate in any significant reactions or sorption (i.e. are conservative). The simple AD model predicts no tangible difference in BTC with a change in diffusivity, whereas the DP model predicts a difference. There are numerous examples of this method in the literature (for example Maloszewski and Zuber (1992) and the chalk experiment of Garnier et al. (1985) which is cited in Maloszewski and Zuber (1990)).

For different solutes the diffusion coefficient will be related to the molecular radius, r_p. This will describe the hydrated molecule and may be only weakly correlated to the atomic mass. [54] It is therefore best to rely on tables of diffusion coefficients for the tracers in question (e.g. diffusion coefficients of common tracers are given in Barker et al. (1995)).

The simplest use of multiple tracers is to simultaneously fit the BTCs of two tracers. An attempt was made to model the effect of varying the ratio of diffusion coefficients of two tracers on the CIs of the parameters t_{cb} and σ for a DP model. The intention was to establish quantitatively to what extent it would be worthwhile selecting the pair of tracers in order to maximise the ratio of diffusion coefficients. An impulse response was modelled and the MC

---

[54] Bodin et al., 2003) cited an expression of the free diffusion coefficient as $D = \frac{k_B T}{6\pi\mu r_p}$, where k_B is the Boltzmann constant, T is the absolute temperature (in $^{\circ}K$), μ is the dynamic viscosity and r_p is the radius of the solute molecules (Einstein, 1956).

method was employed with Gaussian noise of $(m_0/V_m)N(0,0.0015)$. Plotting the CIs of t_{cb} and σ against the ratio of diffusion coefficients gave reducing CIs with increasing ratios. However, depending on the 'base' t_{cb} and σ, the curves were not always smooth. Figure 8.4 gives an example of a typical set of results for one set of 'base' conditions.

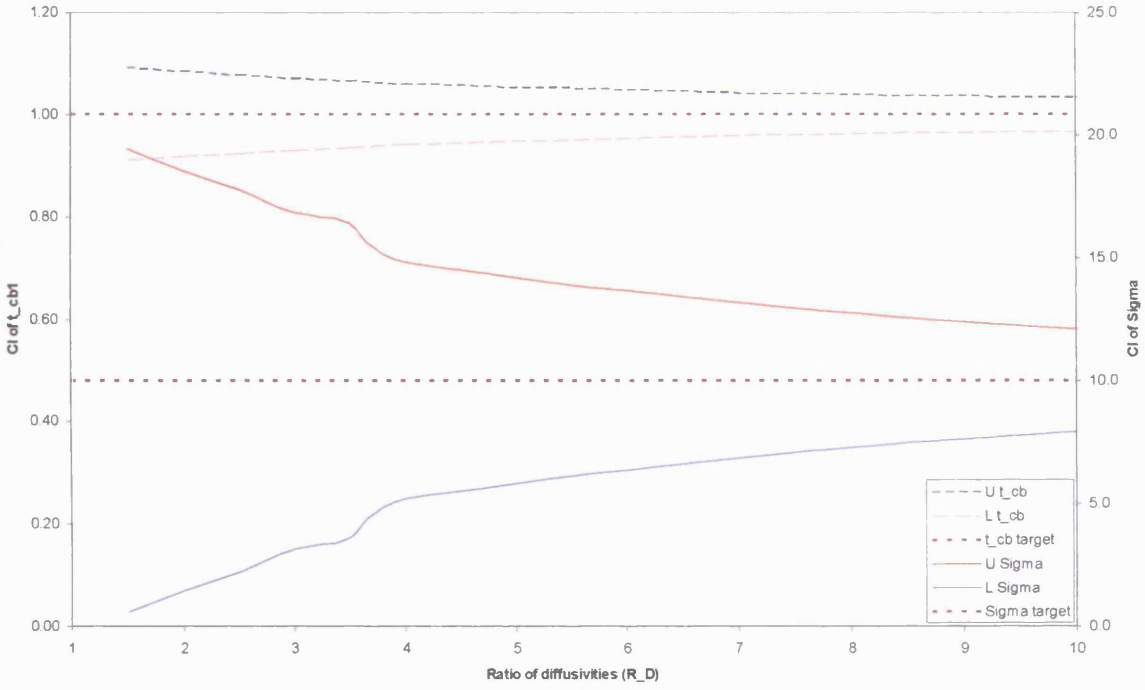


Figure 8.4: CI defined by the upper (U) and lower (L) 95% CI for t_{cb} and σ with a target impulse response BTC with $t_{cb1}=1$ $\theta=0.2$, $q=1$, $z=50$ and $m_0/V_m=1$, therefore $t_a=\dfrac{z\theta}{q(1+\sigma)}=0.909$ and $\sigma=10$. The ratio of diffusivities is $R_D=\dfrac{D_{A2}}{D_{A1}}=\dfrac{t_{cb1}}{t_{cb2}}$.

Figure 8.4 shows an odd jump in CIs at $R_D=3$. Unfortunately, non-elliptical error space confounds the use of CIs established by MC analysis in this way. The findings for other 'base' conditions are therefore not presented here.

Nonetheless, the general recommendation is still clear: if a range of conservative tracers is available and there is no other constraint on the selection of two, choice of tracers with the greatest difference in diffusion coefficients is desirable.

Diffusion affects the amount of Taylor dispersion in a way that may confound diagnosis of a DP process using multiple tracers. This is because, for Taylor dispersion, the predicted spreading of solute is inversely proportional to the diffusion coefficient just as it is for a DP model. The

possibility of Taylor dispersion may need to be eliminated before differences in BTCs are used to diagnose a DP process.

Use of three tracers would be a further diagnostic improvement from the use of two tracers. The third tracer may help to avoid potentially erroneously deducing that one of the tracers must be conservative and the other sorbed.

8.5.3 Advantages of using a top-hat concentration input function

It was shown in Chapters 5 and 6 that a step input function does not unambiguously reveal non-linear effects, whereas a top hat does. If there is any question over whether non-linear sorption may be participating in transport, then usage of step inputs of tracers should be avoided. A more suitable replacement would be a top hat. A top-hat input function further benefits from the fact that the noise can be measured for a prolonged period of (effectively) constant concentration (once the output concentration has approximately approached the input concentration). This will allow a better assessment of the noise model for the system than estimating the noise based on the residuals (since the residuals may include error due to the fact that an inappropriate model has been selected).

The impulse function, although mathematically very important, is not easy to implement in a column experiment (since any response can be created based on the impulse response using the convolution integral). Unless the injection time of solute can be shown to be much less than the advection time (t_A) of the system, it would be more precise to arrange a top-hat input over a known input time, T_p.

8.6 Diagnosis conclusions

The gathering and summarising of relatively simple results here is important for comparative purposes. The exercise helps to assess existing literature, either placing it, criticising it or confirming it. By illustration a few comments are given below on a few deductions that appear in the literature.

An example of why the tables of BTC moments here (e.g. Table 5.4) are useful is an error that is found in Harvey and Gorelick (2000). Although a temporal impulse is input into a system, the moments for an initial spatial impulse are used erroneously, thereby producing a questionable analysis.

The notion that the ADE cannot predict 'early' arrival or long-tails cannot be supported. The AD predicts a finite but small concentration for all z at $t > 0$. Sanchez-Vila and Carrera (2004) demonstrated on a lin-lin axis that the ADE underpredicts late-time concentration if it is fitted to the peak. However, if sufficiently large dispersivities are assumed, long-tails can still be predicted.

A peak that occurs earlier than the time $t_A = z/V$, contrary to what is stated in Rosqvist et al., 2005, is not diagnostic of 'preferential flow' (e.g. of a two-domain DP model).

White et al. (1986) and Rosqvist et al. (2005) used the median time of a BTC that is approximated by a lognormal time distribution to estimate an average solute transport porosity. White et al. (1986) ascribed the substantial overestimate of porosity to transient flow conditions. However this overestimate may have simply been due to the empirical nature of the assumed relationship between median time and porosity. In estimating porosity for an AD system, t_A is a parameter of the fit and by rearrangement (by definition of t_A) gives $t_A = \frac{z}{V} = \frac{z\theta}{q}$, from which θ can be determined. For lognormal median time, writing $\tilde{t} = \frac{z}{\tilde{V}} = \frac{z\tilde{\theta}}{q}$ recovers $\tilde{\theta}$, but it is difficult to say how this relates to θ.

Johnson et al. (1998) calculated t_A for a landfill as approximately two to three years. Since this 'average residence time' is quite short, chemical equilibrium models are justified. There is a potential confusion here. Although t_A is the mean of the BTC it says nothing about the times spent in each block (characterised in a lumped way by t_{cb}), which could be considerably greater than this. Ultimately it is the flushing behaviour of the system that is of greatest interest and if there are large blocks within the material, flushing may be dominantly characterized by very long diffusion times. The characteristic times of a DP system need to be used with care.

The dataset of Becker and Shapiro (2000) at Mirror Lake is amongst the best available from groundwater tracer experiments and is highly intriguing (Figure 8.4). Fitting by eye on a log-log scale produced a 'reasonable' fit to a Br^- recovery curve. This is despite the fact that the fitted model is a DP(infinite) model with a log-log linear gradient of -3/2 and the data have a gradient of -2. Injection was by a dipole. If the experiment had been in a uniform material, a gradient of -4/3 would have been expected. The lack of a -4/3 gradient suggests that there is channelling or fracture flow that makes the interaction between wells more 1D than a dipole would have predicted. There was mixing in the injection well. Over longer time the effect of

having a mixer at the inlet is insignificant (see Appendix 2). There was a similarity in log-log gradients for all three tracers (i.e. the Br^- as well as HDO and PFBA). This was taken to indicate a lack of a diffusive mechanism, since the different tracers had different diffusion coefficients. An AD mechanism was suggested (by deduction). However multi-rate diffusion can produce BTCs that behave differently to the standard DP or DP(infinite) models, giving different log-log gradients.

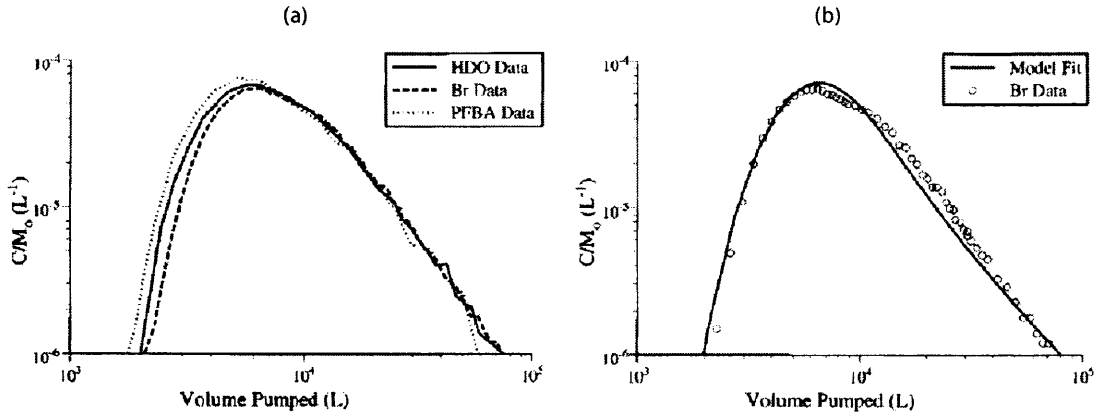
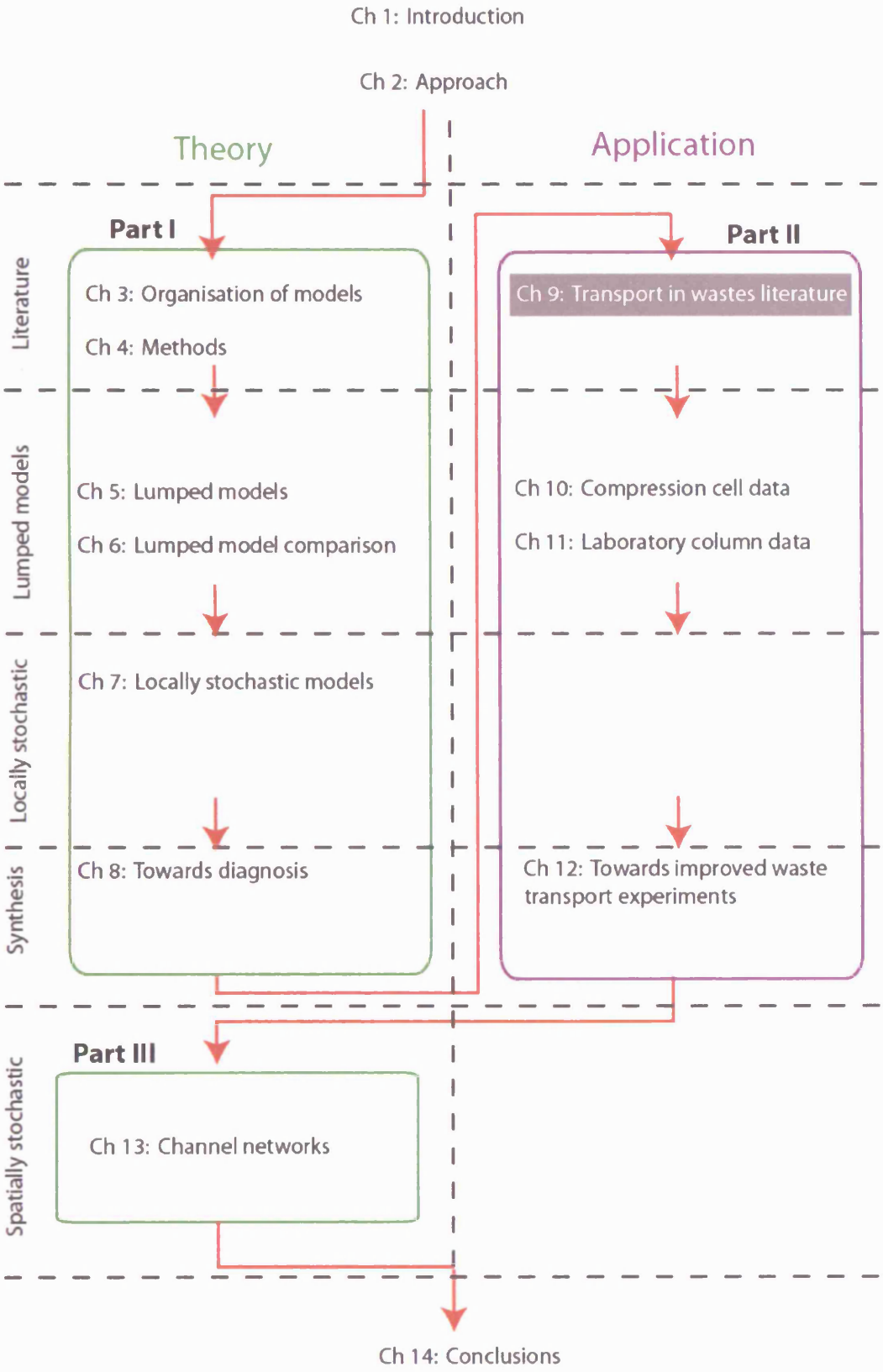


Figure 8.5: Tracer BTCs at Mirror lake (from Becker and Shapiro (2000)) - (a) BTCs for deuterium (HDO), bromide (Br) and pentafluorobenzoic acid (PFBA) ; (b) fit to Br data using an infinite-matrix DP model (log-log -3/2 gradient) with Pe=15, t_a=1620 min, t_{cb}=1480 and borehole mixing time constant 546 mins.

It is worth noting that the expected difference in the curves for a DP(infinite) model would be a vertical offset (not a change in log-log gradient or such a dramatic difference as might be seen on a lin-lin axis, as for example was shown by Garnier et al. (1985)). Interestingly, in a later paper, Becker and Coplen (2001) plotted the same data on a log-lin axis and remarked on how the PFBA data for 'test D' appears to be offset upwards and noted that it could be due to its larger molecular size or lower diffusion rate. It would therefore be worth revisiting this data in the future with multi-rate diffusion models.

So, in conclusion, tools that have been sharpened here allow critical assessment of leading datasets, both in terms of confirmatory findings that may not have appeared when they were published, but also for more critical input. These are now applied to the next part of the thesis (Part II), which is dedicated to grappling with the opposite end of the spectrum of data quality, that of tracing data through wastes. Interaction with the best possible datasets inspires a desire to improve one's own data, but also informs the limit to what can be practically expected.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

PART II: Transport through waste

9 Transport in wastes literature

Part II of this thesis is dedicated to tackling the second aim: *to improve prediction of contaminant flushing from highly heterogeneous materials to the study of transport in wastes.*

This section introduces published flushing data, tracers and models thus far used for developing understanding of transport processes in wastes. This achieves Objective 3: *to assess existing datasets for transport in wastes at a variety of scales and discuss the limits of information content (i.e. how much can really be found out from modelling them).* The contribution of this chapter is a thorough literature review of mass transport in wastes, organised using the heterogeneity scale developed in Chapter 3.

Modern landfills are designed to contain wastes and the leachates that derive from them, in isolation from the external environment. Minismisation of contamination of groundwater is achieved through leachate control. Leachate production is reduced by sealing the landfill with a low permeability 'cap' and surface water drainage to reduce the level of infiltration of rainwater into the waste. Leachates that do arise are collected at the base of the site in order to reduce the hydrostatic pressure on the base lining system. The low permeability liner is designed to restrict flow and transport, however this can never be a complete barrier and may be susceptible to gradual or catastrophic degradation in performance over time.

There is therefore a threat posed to human health and the wider environment from the pollutants held within landfills. A wide range of waste types have been previously placed in landfills and there is an evolving ongoing arrival of new wastes. In the UK alone there is a legacy of around 6,000 closed or unlicensed landfills and around 2,000 licensed ones (Richard Beaven, Personal Communication).

Landfills have a strong potential to represent pollution sources for periods of time measured in centuries (Hall et al., 2004; Krueupelbeck and Ehrig, 1999). There is an urgent need to understand how pollutants are flushed from a broad range of wastes because the landfill potentially forms a 'source' term for pollutant emission into the external environment.

Landfilling of wastes is far from 'sustainable' in relation to most definitions. Of particular concern is the mismatch between the expected time required to operate and then close a landfill (of the order of 50 years) and the period for which it still presents a risk (possibly hundreds of years). Beaven

(2000) defined sustainable landfill as "wastes brought to a stable, non-polluting state within 30-50 years after cessation of landfilling operations". An alternative concept is that of 'passive fail-safe', whereby a landfill is deemed not to be a risk to the external environment even if no management of the site continued.

UK legislation requires that landfill operators must run sites until surrender tests are met, which is ideally when a site no longer poses a threat to the external environment. Various conditions appear in the Environmental Protection Act (1990), the Pollution, Prevention and Control (England and Wales) Regulations (2000) and the Landfill Regulations (England and Wales) (2002). The operator has to demonstrate that a landfill has stabilised (physically, chemically and biologically) to a state in which the undisturbed contents are unlikely to pose a pollution risk. This is done with a risk assessment, which balances the depleting 'source' term of the landfill pollutants against deterioration of the containment systems.

There is no clear consensus about what such criteria for completion should be. It depends on opinion whether tests relating to contaminants should be based on concentration in the leachate, flux of contaminant at the landfill base, or be related to mass still stored. What is considered to be the optimum criterion may be conditioned by the chosen conceptual model of contaminant release and flushing.

The changes in landfill practice that stem from the Landfill Directive (1999) are phased reduction in proportions of biodegradable wastes and treatment prior to deposition and the prohibition of certain types of hazardous wastes (Robinson et al., 2004). These changes tend to reduce biological activity and cause changes in gas and leachate production and composition, and settlement behaviour. Waste volumes are expected to reduce, but with higher contaminant concentrations.

The understanding of transport in wastes is at an important stage. Discussing chemical behaviour, Van der Sloot (2005) discussed the fact that it is now time to move 'inside' the black box of wastes. Van der Sloot (2004) challenged the 'persistent myth' that each waste has a unique behaviour. He suggested that,

> "instead many waste materials behave very consistently and predictably...The leaching behaviour of MSW...is often much more consistent. The leaching behaviour of the fully degraded material may turn out to be similar to that of predominantly inorganic mixed waste".

Powrie and White (2004) took a more cautious line, noting that because of waste heterogeneity[55] deterministic modelling is almost impossible. They suggested that models can be used to carry out calculations based on reasonably representative processes, but that the results cannot be expected to provide an accurate simulation. This is built on work by White et al. (2003, 2004) who noted that highly coupled models have a limited predictive power. They suggested that the value of such landfill models is in providing an effective methodology for organising and assessing complex biochemical datasets, whilst acknowledging the pressing need to make 'quantitative assessments'.

For predictive purposes, single deterministic predictions have limited value and stochastic approaches are essential to account for model and parameter uncertainty. This statement is made irrespective of the noise level of the datasets. Less noisy data will result in a tighter confidence region and will have a tighter predictive envelope (i.e. in the words of van der Sloot (2005), it will be "more predictable").

The approach advocated here for prediction is to develop a stochastic treatment of a pool of simple 'end-member' process models. Models are needed for a wide and changing variety of waste types and conditions, balancing a need to simulate generic processes whilst being tailored to a particular medium. This literature review and the following data analysis (in Chapters 11 and 12) are based primarily on the examples of Municipal Solid Waste (MSW) and the bottom ash from Municipal Solid Waste incinceration residues (MSWi).

Modelling transport in wastes potentially requires inclusion of a large number of simultaneously acting processes, including:

- Advection, mechanical dispersion, diffusive exchange, advective exchange and sorption.
- Gas generation and transport.
- Bio-degradation.
- Geochemical reactions.
- Heat generation and transport.
- Structural (geo-mechanical) change.
- Boundary condition change (input fluids, temperature, placement of waste, overburden and water levels).

[55] This is another example of where the term heterogeneity needs to be interpreted carefully. Here it is taken to infer noisy data, but it could have implied a lack of REV for example.

There are several models which incorporate representations of some of these processes simultaneously. There may be cases in which only a small number of the processes are operating (or dominating), in which case simpler models can be tested.

During organic degradation (which occurs at a relatively early time in terms of the timescales of interest for inorganic flushing) the movement and presence of moisture is essential to the removal of solutes, but also to the movement of nutrients and water for hydrolysis and other reactions involved in anaerobic degradation. Transport may be a central factor in the larger coupled biological and mechanical system. Later on, when the organic chemistry has stabilised, the transport of solute is more important with reference to the flushing of inorganics.

What is needed is organisation and appraisal of the wide range of models that have already been used. This chapter seeks to achieve this for transport models by applying the organisational structure proposed in Chapter 3. Following the current chapter two different experimental datasets are examined and modelled in Chapters 10 and 11. The lessons learned from this experience are gathered in Chapter 12 in which recommendations are made for improved future experiments and modelling development.

9.1 *Landfill contaminant flushing data*

Good landfill field contaminant flushing datasets are very hard to come by. This is primarily because the uncertainties surrounding field datasets can be substantial. In interpreting field data there are many factors to take into account, such as age, waste type(s), cell-layout, drainage, monitoring set-up, climate, cover and gas. It is difficult to control the boundary conditions, so the onus is more on monitoring them, however such data can be scarce.

Waste data can be complicated by a number of further factors:

- Precipitation, freezing, evaporation.
- Continued waste placement, irrigation or other management changes.
- High strength leachates (high TDS).
- Mixing in the extraction system.
- Disturbance by other hydraulic structures (e.g. roadways, wells, gas-abstraction infrastructure, basal drainage layers).

The combined effect of these issues is that tracer and flushing datasets tend to be characterised by high noise, poor boundary condition control and poor control of other process variables. Robinson (1995) noted that any (historically collected) field data that are available should be treated with great uncertainty.

Historically, measurements of landfill flushing appear to have been made with conceptual models in mind that relate flushing to the liquid to solid ratio (LS), rather than time.[56] This is a particular difficulty for Dual Porosity (DP) type models for which time is a more pertinent parameter than cumulative flow.

The implications for improvement of future experiments and monitoring programmes are discussed in Chapter 12 where this principle of tackling each issue relating to lack of control, noise and uncertainty is pursued. Although wastes pose a considerable challenge, a careful reductionist approach to their measurement and modelling is potentially able to make considerable inroads.

9.2 *Tracers in wastes*

9.2.1 Uses of tracers in wastes

Tracers are potentially of immense value to understanding flushing in wastes. Two primary uses of tracers in waste can be identified in the literature. Firstly, there has been conventional artificial tracing, by addition of tracers at the surface of landfills by irrigation. Secondly, tracers have been used as 'signatures', for example, isotopic fractionation may be indicative of biochemical action or evaporation, indicate passage through the top-cover or reflect recent rainfall. Certain key species have been used to indicate transport from a waste mass into the external environment (for example, the use of CFCs).[57]

9.2.2 Types of tracers for wastes

Blakey et al. (1998) reviewed the selection and use of tracers for investigation of transport in wastes. A broad spectrum of possible tracers was considered - radioactive tracers, dyes, CFCs,

[56] LS is defined as the total cumulative liquid (leachate) that has passed through the waste, divided by the dry mass. The purpose of using LS is to enable comparison of the flushing states of different amounts of waste (with the presumption of similar starting concentrations of contaminants).

[57] This was an unpublished demonstration of concept by T. Atkinson and Zhao (2000), at UCL: "Using Chlorofluorocarbons to Detect and Monitor Landfill Leachate Plumes in Groundwater and Rivers", as part of the EPSRC Waste and Pollution Management Programme.

biological (bacteriological and spores) and inorganic chemicals. They documented first-hand experience of only two, tritium and lithium chloride, both of which proved to be good tracers.

In general, tracing in wastes is a difficult task. The wastes themselves are heterogeneous mixtures of materials and have leachates that have a high TDS comprised of many different chemicals. This means that the number of possible reactions is high, the background concentrations of specific species or properties (e.g. EC, fluorescence) can be high and analysis can be difficult in the presence of high strength leachates.

The reported tracer experiments to date are summarised in Table 9.1.

Table 9.1: Tracers used in landfills

Type	Species	Location	Author(s)	Waste type	Duration	Cell details	Comments
Inorganic	LiBr	Filorna, Sweden	Öman and Rosqvist, 1999;	MSW	80 days	Pilot-scale	
	LiBr	Location unspecified, Sweden	Rosqvist and Bendz, 1999; Bendz and Singh, 1999;	Shredded, degraded MSW	8 hours	Field-column	Built into an existing landfill
	LiBr	Pitsea, UK	Beaven et al., 2003;, Woodman et al., 2005;	DANO-MSW	146 days	Lysimeter	Difficulties analysing Br
	LiBr	Brogbrough, UK	Beaven et al., 2001;	MSW	21 months	Test-cells	BTC Not clear, apparently bi-modal.
	LiCl	W. Yorkshire,UK	Blakey et al., 1998 cites Blakey, 1996;	MSW	262 days	Recirculation cell	
	KI	Lostorf, Switzerland	Johnson et al., 1998;	MSWi bottom ash	22 months	Landfill	Only qualitative treatment
	Sr	Teuftal, Switzerland	Ludwig et al., 2000;	MSWi bottom ash		Landfill	Natural tracer (rainwater)
	H_2O	Teuftal, Switzerland	Ludwig et al., 2000;	MSWi bottom ash		Landfill	Dilution (effect on EC)
Isotope	^{2}H	Southampton, UK	Beaven et al. (unpublished);	Digested MSW	1000 hours	Laboratory-cell (recirculating)	
	^{3}H	Medmenham, UK	Blakey et al., 1998 cites Blakey, 1982;	MSW	18 months	$8m^3$	
	^{18}O-^{2}H	Technical University of Brunswick, Germany	Maloszewski et al., 1995;	MSW	3 years	Lysimeters	Mixing model and simple AD model
	^{18}O-^{16}O	Lostorf,	Johnson et al.,	MSWi	22	Landfill	Mixing model to

Type	Species	Location	Author(s)	Waste type	Duration	Cell details	Comments
		Switzerland	1998;	bottom ash	months		determine proportion of direct throughflow from rainfall
	^{10}B-^{11}B	Riet, Switzerland	Hoen et al., 2000;	MSW	Spot samples	Landfill	$^{10}B/^{11}B$ ratio was used to identify signatures from different cells in the landfill
	^{82}Br	Riet, Switzerland	Hoen et al., 2000;	MSW	Rapid	Landfill	Single borehole dilution tests (γ-ray probe) emissions measured
Dye	Brilliant Blue	Sweden	Rosqvist and Bendz, 1999;	Shredded, degraded MSW	Hours	Field-column	Built into an existing landfill
	Brilliant Blue	Southampton, UK	Beaven et al., 2005;	Chopped MSW	50 days	Laboratory-column	Sorption study by Jing-Lin (2005)
	Uranine	Breitenau, Austria	Döberl et al., 2003;	MSW	Approx. 1 year	Landfill	Multi-model arrival, no attempt at modelling
	Pyranin	Lostorf, Switzerland	Johnson et al., 1998;	MSWi bottom ash	Approx 1 year	Landfill	CXTFIT model of arrival (0.4% recovery)
	Fluoresce in-Na	Lostorf, Switzerland	Johnson et al., 1998;	MSWi bottom ash	Approx 1 year	Landfill	Only qualitative findings (non-typical BTCs and no information on sorption)
Gas (PGTT)	He-difluoro methane	Delaware, USA	Han et al., 2006;	MSW	800 mins	Laboratory-scale and landfill-scale	Moisture content estimates by comparison of retardation due to affinity with water of two different gas tracers. Various problems and inaccuracies under low saturation conditions

*Unpublished data, by Beaven and co-workers at the University of Southampton

9.3 *Models for transport in wastes*

9.3.1 Heterogeneity and preferential flow

Wastes, such as MSW, are evidently 'heterogeneous'. They are comprised of a range of differently shaped, sized and coloured objects that look fairly large to the eye, in amongst more 'soil-like' brown mulch. What this heterogeneity means for waste measurement and modelling is potentially very significant.

Figure 9.1: Untreated, aged MSW (photograph from Richard Beaven).

There needs to be clarity about the term 'heterogeneity', because it only has meaning with respect to a given scale. The visual scale however, does not matter as much as the scale of physical mixing.

If the reference scale is the domain-size, then 'highly heterogeneous' could mean 'not amenable to continuum representation', as the properties in question may continue to vary beyond the scale of the domain. However, the reference scale could be of the order of tens of centimetres, which is certainly heterogeneous to the eye, but would not necessarily invalidate continuum approaches. Continuum AD models can be derived from spatial heterogeneity (variable permeability), but only if the domain size is sufficiently large (Gelhar, 1993). Continuum DP representations rely on plenty of local flow variation to cause diffusive disequilibrium on one scale, but at the same time to also

present a similar pattern on a larger scale at all points. The term 'heterogeneity' and 'preferred pathways' are occasionally used synonymously, which is potentially confusing.

Zacharof and Butler (2004a) inferred that 'heterogeneous' implies that the system is not amenable to deterministic representations and are presumably referring to the level of noise in the data (which perhaps they imagine comes about through measuring at a smaller scale than the heterogeneous length scales.) They argued convincingly that since only crudely defined pdfs of parameters can be found, the modelling treatment should be stochastic in order to accommodate such uncertainty. Yet, their model appears to imply a different understanding of heterogeneity. They define a 'statistical homogeneity' (the same as that defined by Greenkorn and Kessler (1969)) and use what is effectively a continuum model in which their 'heterogeneity parameter' is equivalent to a dispersion coefficient. This understanding of heterogeneity does not necessarily infer noisy data or broadly spread parameter pdfs and implies that the heterogeneity scale is considerably smaller than the measurement scale (and certainly the domain scale).

Field observation has provided evidence of preferential flow, leading a number of authors to present dual-domain representations of wastes. A number of studies provide evidence of dual-domain behaviour (Bengtsson et al., 1994; Bendz and Singh (1999); Blakey, 1982; Blight et al., 1992; Ehrig, 1983; Harris, 1979; Holmes, 1983; Stegmann and Ehrig, 1989; Uguccioni and Zeiss, 1997). Evidence of channels created by bridging structures created by larger waste items is given by Wall and Zeiss (1995). The distribution of moisture has been shown to be spatially variable as well as temporally variable (Bendz et al., 1997). Yuen et al. (2000) reviewed attempts at measuring the spatial distribution of moisture showing that there are difficulties with using tensiometers and gypsum blocks, but that the neutron probe offers the possibility of at least measuring relative changes in moisture content.

Hudson et al. (1999, 2001) showed considerable variability in K with increasing applied stress (simulating increased depth) in the Pitsea compression-cell. Gawande et al. (2005) used moisture sensors to infer a range of Ks by adjusting K in the 3DFEMWATER code until moisture arrival from injection wells was correctly predicted at the sensors. Data from the same experiment was interpreted by Jain et al. (2006), using a borehole permeameter test. They also reviewed a wide range of studies that measured K, showing a very large range in values.

There are no visual reconstructions of the 3D flow structure at representative scale in wastes. Where there have been cuts made into a waste 'face' the structure can be examined analogously to

rock fracture patterns at an exposure. There have been attempts to freeze and section wastes (Beaven, personal communication), but as yet there is little quantitative data. There is a strong need to characterise structure and flow patterns by direct observation.

In some instances the geometry is well-known (or at least the geometry of the solids at the time of placement). Ludwig et al. (2000) documented 0.5 m^3 cubes and also pellets of cement-stabilised MSWi Air Pollution Control (APC) residues. Of all the wastes, MSWi wastes hold out the prospect of being most amenable to measurement and characterisation. MSWi wastes are thought to hold remarkable similarities in composition and structure and most closely resemble a classical granular porous medium, at least at the time of placement. The most likely conceptual model is therefore of AD. Having said that, Johnson et al. (1998) observed evidence of preferential flow and used the MACRO model to model this effect. They did not suggest a conceptual model to explain this, but perhaps there are two pore scales (granules and micropores) or even self-organised flow structures. The possibility of cementation (creating larger aggregates) or flow organisation (creating larger dead-zones) has not been eliminated.

Somewhat less conclusive evidence of preferential flow comes from the BTCs generated by tracer tests, as neither 'early' breakthrough nor observation of a 'long-tail' are diagnostic (Chapter 3). Intermittent flow tests (Beaven and Hudson, 2002; Beaven et al., 2003) reveal existence of a dual-porosity transfer process but do not infer that it is dominant.

The structure that the fluid-flow 'sees' through this depends on flow-path organisation, which to some extent also depends on water saturation, gas level and flow rates. It would be expected that the freshest, youngest and least compressed wastes near the top of a landfill would have the biggest flow-paths with more compressed waste towards the base. As wastes degrade it might also be expected for them to compress and become more homogeneous with fewer large pathways.

The role of gas potentially adds considerable complexity. Dynamic path-switching has been observed for unsaturated flow (Glass et al., 2002). Hudson (personal communication) notes that this may be an important process for a timescale of the order of minutes and hours in a lysimeter. Because of the loose nature of MSWs there also exists a possibility that there is a fluid-structural interaction, perhaps with particles in the waste itself being able to move, perhaps acting as valves moved by growing bubbles of gas. Therefore, by several possible mechanisms, gas may lead to flow instability.

Direct observation might be able to provide quantitative data, for example, to better constrain the DP model, channel shape spacing, immobile block size distributions and potentially demonstrate existence or lack of existence of hypothesised effects such as 'valves' and flow switching.

Conceptual models of MSW structure have been suggested. The difficulty with presenting a single conceptual model is the extent to which it is presumed to be representative of a large number of sites. There has not yet been a systematic attempt to categorise wastes in terms of their internal structure, nor make a linkage to transport.

Bendz et al. (1998) conceptualised three heterogeneity scales; the 'lens' scale of 1 m horizontally (a refuse element), the truck load scale and the landfill scale. They conceptualised sub-parallel layers of squashed plastic bags with tears and gaps between them which allow vertical flow (Figure 9.1). In their flow model for 22-year-old waste, landfill gas formed a fairly passive continuous phase that displaces what would otherwise be air. This may not be applicable to early-time flow situations where liquid-gas interaction is more important.

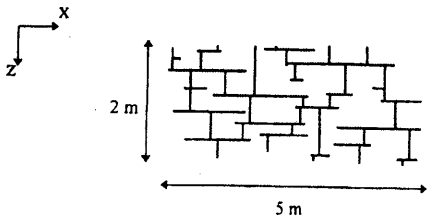


Figure 9.2: Conceptual model of sub-parallel layers connected by vertical flows, from Bendz et al. (1998:2964)

Fellner et al. (2003) hypothesised that waste-structure dominates flow patterns. They emphasised that this is different to soils, which are often used as an analogy. Impermeable objects in wastes give a greater extent of perched water and horizontal flow around these blockages in contrast to 'predominantly vertical' flow in soils. Bendz et al. (1998) suggested that this may be a relatively minor effect, but ultimately field evidence is required to verify this. Figure 9.2 conceptualises the effect of perched water and less-flushed regions. Fellner et al. (2003) conjectured that there are lots of 'horizontal feeders' to a channel, with more matrix flow near the top and more channel flow near the bottom as matrix water enters the channel.

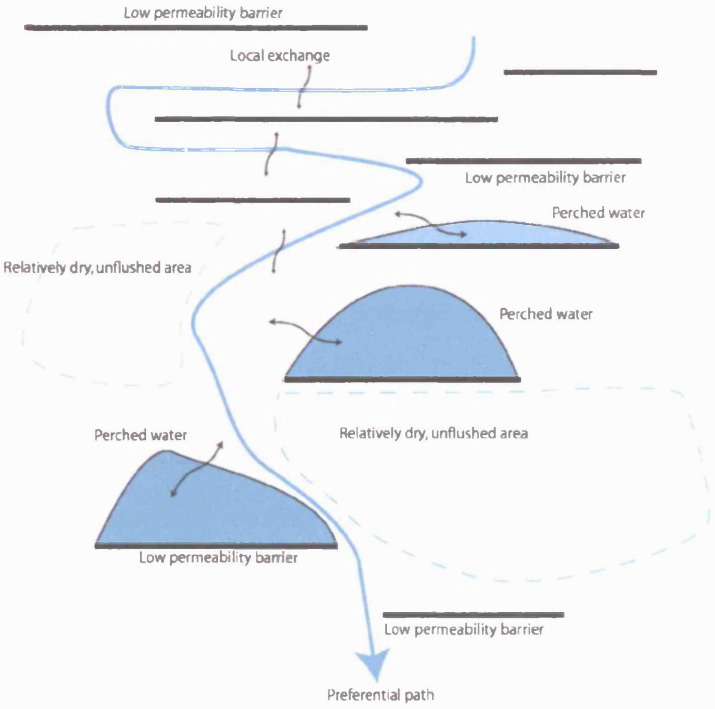


Figure 9.3: Conceptual model indicating key features for flows through wastes

Döberl et al. (2003) equated the notion of preferred paths to heterogeneity[58]. They also make a convincing argument about how to interpret concentration data from a DP system, pointing out that,

> "the substance [contaminant] load [at the base of the landfill] represents the surrounding area of the preferential flow paths only. Hence, the substance concentration, which can be measured at the bottom drainage, represents only a very small percentage of the total landfill volume. This indicates that the assessment of the future emission potential of MSW landfills should not be restricted to substance concentration only."

They touched on a very important facet of DP system behaviour - that the concentration measured may be considerably below what is stored in the matrix and is not a good measure of the total mass balance[59]. The distinction between the interpretations of heterogeneity has been emphasised here

[58] They created a Hydrus2D discrete channel model, in which the matrix is represented by a 2D continuum adjacent to a channel.

[59] Although simple, this conceptual point can lead to inaccurate estimates of contaminant mass left in a system and under a changed hydraulic regime concentrations that could emerge could be suddenly considerably greater. An equivalent underestimate can happen for pump and treat systems when pumping is stopped because the pumped concentration is low.

because modelling heterogeneity in a DP system may require representations such as a range of block sizes or a fully spatially stochastic approach.

The choice of conceptual model can influence wider issues, including that of how to decide on completion criteria, in order to design an experimental programme, to monitor, remediate or parameterise.

There have been numerous attempts to capture heterogeneity, for example:

- 'Giving up' and adopting a single mixing-cell model.
- Adopting a continuum model with a 'heterogeneity parameter' (Zacharof and Butler, 2004a) and treating all parameters stochastically.
- Adopting a conventional AD model.
- Using a classical stochastic approach to test the effect of distributions of uncertain parameters on a continuum model in order to respond to noisy data (Zacharof and Butler, 2004b).
- Using a stochastically-generated permeability field (McCreanor and Reinhart, 2000).
- Conceptualising increasing levels of heterogeneity from lumped, through locally stochastic to fully spatially stochastic (this study, Part III).
- Obtaining parameter uncertainties by calibration, not 'guesswork' (Chapter 5).

The following section reviews the transport codes for wastes. It is shown that there are a large number of potential processes to incorporate. How heterogeneity is represented is a key question which must be asked of every representation.

9.3.2 Codes for modelling transport in wastes

9.3.2.1 Introduction

A summary of different approaches to waste modelling up to 1997 was given by El-Fadel et al. (1997). They described various ways of incorporating leachate generation and degradational terms in conjunction with a transport representation. For the models covered by El-Fadel et al. (1997), transport is usually based on an EPM continuum assumption, which varies from simple water balance to AD. They noted that, up to the point of their paper, focus was primarily on the liquid phase and there was scope for work on the multiphase nature (solid-liquid-gas interactions). What is also conspicuously absent at this stage is the use of models that explicitly incorporate preferential flow and exchange (either by advective flow or diffusion) with less mobile zones.

The balance of model simplicity with the known complexity of the system is not always directly referred to, though Zacharof and Butler (2004a) consciously struck a balance by adopting a parsimonious process model akin to the level of abstraction preferred in this work. As crucial as striking a balance of simplicity is getting the essence of the core processes correct, or covering a range of possibilities if none can be determined with certainty. Zacharof and Butler (2004a) are limited by fixing on an AD model and omitting representation of preferred pathways or transfer terms.

A further review of codes is presented here, organised in order of increasing levels of incorporation of heterogeneity. Thereafter, models that sit outside this categorisation are considered.

9.3.2.2 Continuum models for transport in wastes

9.3.2.2.1 Advection-Dispersion models for transport in wastes

Flow through some wastes, particularly incinerated, treated, sorted or shredded wastes, might reasonably be approximated by inter-granular flow. Other wastes may have more of a network of interconnecting pathways.

Since a series of coupled mixers can result in AD-like behaviour, it is worth noting that models which do this may inherently simulate the AD process, at least incipiently (for example the stack of well-mixed cells in Yildiz et al. (2004)).

Due to concern over increased advective fluxes to the external environment due to higher 'leachate table' levels, many modern landfills are kept primarily unsaturated with saturation only at the deepest levels.

The 'saturated' AD model has appeared under various guises. The AD (with an apparent dispersivity that can include diffusive transfer) was used by Maloszewski et al. (1995), CXTFIT (a saturated AD model) was used by Johnson et al. (1998) and Rosqvist and Bendz (1999) used a stochastic lognormal time-distribution model. With a simple rearrangement it can be seen that the hydraulic model of Zacharof (2001) is identical to the well-known lognormal stochastic-advective model of Jury and Roth (1990).

For unsaturated flow if the water flux is variable at the inlet to the system, the water contents, permeabilities and flow rates will vary with depth. The unsaturated system is potentially very different to the saturated one in which the hydraulic diffusion time for a fully saturated system in response to a change in flow rate is very fast. The hydrological study of Bendz et al. (1997) showed seasonal recoveries and recessions in water contents and that it can take several years for freshly deposited waste to achieve 'field capacity'. Wastes that are not artificially irrigated experience considerably transient degrees of saturation.

Johnson et al. (2001) used Hydrus5 to model unsaturated flow through MSWi bottom ash by Richards' equation, and AD thereafter, but found that it failed to capture the dynamics that were leading to fast responses to rainfall events.

Straub and Lynch (1982) and later Demetracopoulos et al. (1986) solved Richards' equation for flow and then applied the AD equation in conjunction an empirical source term (a modified MIM transfer) to model inorganic flushing.

There is a possibility that AD representations may be reasonable even though preferential channels and immobile zones are known to exist. Chapter 10 contains a discussion of whether the AD model may still reasonably represent lysimeter data, despite the demonstrated existence of immobile zones.

9.3.2.2.2 Dual Porosity models for transport in wastes

Woodman et al. (2005) described using diffusive models for waste transport and flushing assuming a constant flow. The considerable simplification of assuming constant flow is worth further discussion. Such models are commonly used in the soils literature (e.g. Griffioen et al. (1998) compared a large number of results where the simple two-domain model has been applied to unsaturated media and van Genuchten and Dalton (1986) examined two-region models for use in aggregated soils). Of course, the assumption of hydraulic equilibrium does not necessarily imply that the system is fully saturated. A column experiment (or well controlled field flushing system), by having a constant inlet water flux, could produce conditions under which there is hydraulic equilibrium. For example, this could be either in an unsaturated state with a negative matric potential, or in the presence of a positive gas pressure (e.g. the Pitsea cell experiment, described in Chapter 10).

A feature of having dual-domain type models is that the contribution to the exchange term due to physical processes other than diffusion may not be obviously distinguishable. The effect of sorption kinetics, diffusive transfer and advective transfer may potentially be combined into a single 'apparent' parameter. An important reason for investigating the diffusive term is that it is potentially the slowest form of transfer. There is good evidence that some datasets require the hydraulic transfer term to account for the breakthrough following rapid through-flow after short-term rainfall events. However, it is worth first asking the question as to what proportion of the mass flushed is affected by getting this short-term feature correct and secondly whether there is also a diffusive term that affects longer term flushing (which ultimately is the most important issue)?

The 'end-member' models used in this thesis (and introduced in Chapter 5) potentially concern the fraction of waste which is not hydraulically accessed and is only flushed diffusively. A first order representation (Bendz and Singh, 1999) may not necessarily give a conservative estimate and it is worth exploring the options for how this fraction may be flushed.

The reason given by Johnson et al. (1998) for how 'events' of more diluted solute arriving at the base are observed in MSWi bottom ash was that some of the 'clean' liquid in the channels that is not taken into the matrix gets down rapidly to the base and blends with leachate from the matrix. A weakness of this model is that the entire 'microporous' zone ('matrix') is assumed to be described by Richards' equation. Because there is a very strong possibility that in reality there are large 'dead' areas which do not convect, this model potentially over-predicts flushing rates.

Bendz et al. (1998) presented a model for water movement, followed by a paper built on the same model but with additional consideration of solute transport (Bendz and Singh, 1999). Interestingly, in common with the approach here, they avoided multiphase flows or other coupling, adsorption or dissolution terms and consider only water flow and conservative solute transport. Water movement is imagined to be film flow and is modelled by a kinematic wave in the mobile zone along the same lines as for the MACRO model. Capillary forces are assumed to be negligible in the channel. Unlike the DK aspect of MACRO, the matrix is considered to be immobile except for local advection to and from the channel. This imbibition to the matrix is determined by mass-balance in the preferential flow zone (a positive imbalance between inflow and outflow must go into the matrix). Drainage is modelled by a slower first order term (which also therefore decays as the total matrix water is diminished). Figure 9.3 summarises the conceptual model used by Bendz et al. (1998). This model was calibrated successfully against transient water flow in an experimental cell with waste.

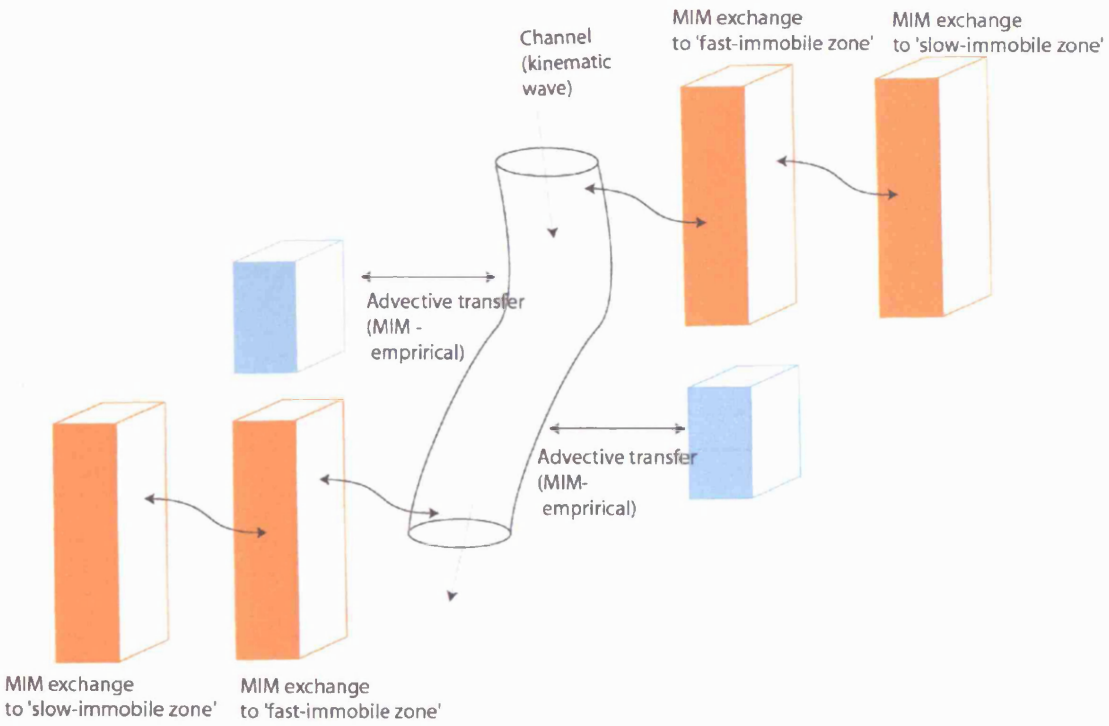


Figure 9.4: Two immobile diffusive zones and one immobile advective zone, based on the model of Bendz et al. (1998)

The solute transport model includes an advective transfer due to the transfer between zones (with the first order drainage term replaced by the same simple mass balance that is used for imbibition). There is also a three-domain diffusive DP system, with a mobile domain coupled to an immobile-fast domain (to represent the more degraded near-channel material) which in turn is coupled to a deeper immobile-slow domain. These are all modelled as first-order (MIM) diffusive transfers.

The MIM return term for the advective component is not mentioned in Bendz and Singh (1999), so is perhaps not included anymore. If unsteady-state conditions occur, fitting was improved by increasing the fraction of matrix allowed to undergo first-order exchange. This is physically justified because the additional advective exchange could be expected to expose a greater volume of solute to subsequent diffusive exchange. A natural practical outcome from this suggestion is that pulsed flushing might be achieved more rapidly than continuous flushing. Interestingly, with a diffusive-only conceptual model pulsing does not necessarily enhance the flushing rate, which would appear to be counter-intuitive (Harvey et al., 1994). The advective store is not coupled to the diffusive store in the model, so there is a possibility that this finding is just a quirk of calibration.

The use of the fraction of 'fast' immobile store, f, that is active in transfer is difficult to visualise physically, as there is no account as to what happens to the proportion that is left ($1-f$).

The fits to experimental data are reasonable, although the early time is poorly represented due to the sharp-shock arrival front which is a one of the weaknesses of having a MIM-only model (Chapter 3). The tail for the steady-state flow model appears slightly over-predicted and the transient fit looks very encouraging, but it had only been flushed to about 20% of the peak concentration (not a very testing case).

Rosqvist and Bendz (1999) were self-critical and note the arbitrariness of several of their assumptions. They note that the lack of spatial variation is potentially a weakness, the MIM parameters would be difficult to determine by physical measurement and the transfer coefficients are likely to have lumped in the effects of mechanical mixing and sorption[60]. Their finding about increased diffusional exchange in the presence of advective exchange is dependent upon their conceptual model being an accurate representation, which they themselves concede is something far from proven.

9.3.2.3 Locally stochastic models for transport in wastes

Johnson et al., 2001 used the programme MACRO (Javis, 1991), which employs a 'kinematic wave' description for channel flow. This models unsaturated hydraulic conductivity as a power function of saturation, as first postulated by Beven and Germann (1981) when unsaturated and just uses mass-balance when flow is fully saturated.

For the relatively low water contents expected in landfills, this model will tend to only direct flow out of the preferred zone. Exchange in the opposite direction can occur if the micropores are saturated (either at the surface or at a layer where the micropore conductivity is reduced), whereby all the excess water in a layer is immediately transferred to the macropores. In near-saturated systems this can have an important role (van den Daele, 2005). This could be modelled as occurring in the instance of perched water above major impermeable objects. Such a conceptual model would potentially also predict a poorly swept zone beneath such an object with a lower water content and therefore not be likely to supply any outward advective flow (and have to rely on diffusion for flushing).

[60] Data from Stegemann et al. (2006) as well as in Reinhart's PhD thesis (Reinhart, 1989) shows that lithium is effectively not sorbed to MSW (but may be slightly sorbed to more aged wastes).

MACRO solves Richards' equation for the 1D vertical 'matrix' with a source term for exchange with preferential pathways 'macropores'. Since macropore flow is modelled as gravity-driven, flow pressure is not required to be calculated. Exchange between the zones is modelled by a semi-empirical MIM exchange term. The term is based on the rectangular-slab geometry combined with an empirical linkage to saturation (Gerke and van Genutchen, 1993). Figure 9.5 illustrates the conceptual model behind MACRO.

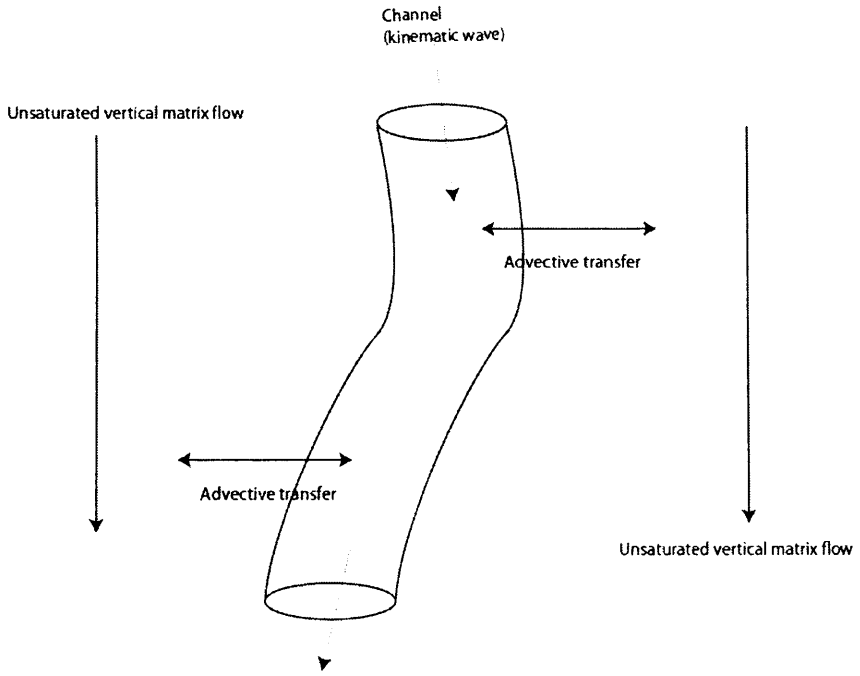


Figure 9.5: MACRO-type dual permeability.

9.3.2.4 Spatially stochastic models for transport in wastes

McCreanor and Reinhart (2000) used the UGGS code SUTRA on an uncorrelated stochastically variable permeability grid. The permeability values were randomly selected from either a normal or exponentially increasing or decreasing functions.

Spatially stochastically generated systems can potentially produce models that are at a sub-continuum scale. Such representations are interesting, but potentially very difficult (or impossible) to calibrate. Also, flow paths may be sub-continuum, but it is possible that the matrix is not.

Other than McCreanor and Reinhart (2000), there are very few reported attempts to represent heterogeneity in a spatially stochastic manner. In the soils and groundwater literature this has proved to be a very important line of investigation, but one which is dependent on reasonable geostatistical data, or physical conjecture. Investigating the latter may reveal classes of behaviours that arise due to particular assumptions in a given representation. In this light, one such conjectured representation is examined in Part III.

9.3.2.5 Empirical models for transport in wastes

The simplest and earliest model for flushing of wastes is the 'continuously mixed reactor' model (Figure 9.6). The waste is imagined (unrealistically) to be contained in a tank which is mixed.

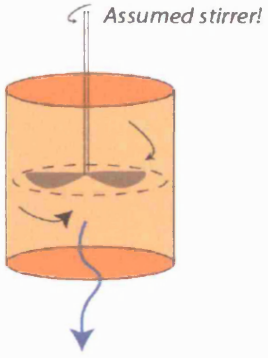


Figure 9.6: Schematic of the 'continuously mixed reactor model'.

The mixing model remains the dominant empirical model. It is an end-member for 'perfect' conditions for biological and chemical reactions to occur with all the other reactants present, rather than them existing somewhere else in the landfill.[61] A more interesting question is how and why an 'exponential model' might be an appropriate end-member. As is shown in Chapter 5, the AD, DP and MIM models all tend to exponentials at late-time (i.e. have log-lin straight gradients). To be

[61] Beaven (2000) suggested a different end-member flushing model, called 'fill and draw', in which the landfill is periodically filled and emptied. Assuming a DP system which achieves equilibrium following a fill, the immobile concentration after N cycles will be $C_{im}^{N} = C_{im}^{\text{initial}} \left(\frac{\theta_{im}}{\theta} \right)^{N}$.

used in this sense the parameterisation of the model would have to be related to physically-based concepts.

Gönüllü (1994a) presented a coupled model for inorganics. They extend the 'well-mixed reactor' as developed by Straub and Lynch (1982) by adding in the source term that Straub and Lynch (1982) used in the AD model. In a second paper Gönüllü (1994b) applied a similar approach based on the 'well-mixed reactor' but coupled it with a simple biological model.

9.3.2.6 Other features for modelling transport in wastes

Amongst the other features that have been modelled in landfills are 2D flow toward horizontal drains (Lu, 1996), and the effect of adding cells at different times (Yildiz et al., 2004).

A very significant class of models is the highly coupled models, which variously combines geo-mechanical properties, degradation, gas generation and transport. For example, El-Fadel et al. (1996) described a coupled model which is purely for gas transport including coupled terms for biokinetics, heat and advection. The effect of waste compression is included in the LDat model (White et al., 2004), which used the empirical relationship found by Hudson et al. (2004). LDat also incorporates anaerobic biodegradation. The transport of the liquid phase in LDat is by Darcian flow though an EPM. Mechanical dispersion could be included by assuming a distribution of permeability of different volumes at the grid-cell level.

The hydro-bio-mechanical flow model of McDougall and Silver (2005) solves classical unsaturated flow equations using the empirical moisture retention curve and conductivity relationships given by van Genutchten (1980). They cite their data in support of this (Kazimoglu et al., 2006). McDougall et al. (2004) likens MSW waste to a 'two-tier' pore structure (i.e. aggregates with micropores), using data from the Lyndhurst landfill (Victoria, Australia) to support the analogy. Cellulose, as a major component of the organic fraction, has systems of large and smaller capillaries analogous to soil micropores and macropores. Codes such as the one by McDougall and Silver (2005) are very useful for looking into hitherto unknown coupling effects that arise from permutations of processes.

The multiphase nature of flow with landfill gas present is difficult to conceptualise and model. If the focus is on solute transport it is tempting to assume that the gas simply replaces air in a well-connected gas-phase that coexists with the liquid and solid phases (Bendz and Singh, 1999). The gas might be presumed to vent out with little interaction with the solid or liquid phases. It would be

useful to establish whether this becomes a reasonable assumption at certain waste densities, water contents and gas production rates. However, at high gas production rates, in the presence of a high water content, the gas-liquid interaction is presumed to be highly complex. Alternatively the gas phase might be assumed to be either stationary or in quasi-equilibrium and effectively representing a solid in the path of the liquid flow. This sort of assumption might be useful to make in the context of short-term tests relative to the degassing of the waste.

9.3.3 Summary of transport in wastes

A selection of the latest of the models discussed here is shown in Table 9.2.

Table 9.2: Summary of recent waste models.

Key: DP - Diffusive dual porosity, ADP - Advective dual porosity (due to transient saturations), Bio - Biologically mediated organic reactions, GeoChem - Geochemical reactions, Mech - Geo-Mechanical effects, UAD - Unsaturated advection-dispersion , AD - Saturated advection-dispersion.

Type	Name	Author	Abilities	Limitations
DP (&AD)	DP-PULSE, DP1D, LT	Barker et al. (2000). Dispersion added in this thesis (Appendix 1)	Variable flow rate, reverse flow, different geometries, distributions of block sizes. Successfully fitted lysimeter and field data	Does not model: Bio, Mech, ADP transfers, GeoChem
DP & ADP	-	Bendz and Singh (1999)	Successfully fitted lysimeter data (exception of early time)	Does not model: Bio, Mech, GeoChem. ADP transfer is not fully physically based. DP transfer is first order. Two immobile zones for DP difficult to evaluate experimentally. Highly parameterised
ADP	MACRO	Johnson et al. (2001)	Successfully fitted field MSWi bottom ash data following rainfall events	Does not model: Bio, Mech, GeoChem. ADP transfer is empirical. Highly parameterised
UAD	HYDRUS5	Johnson et al. (2001)	Richards' equation. Failed to model post-rainfall transients	Does not model: Bio, Mech, GeoChem. No DP or ADP transfers

Type	Name	Author	Abilities	Limitations
Discrete fracture	HYDRUS2D	Fellner et al. (2003)	Models the supposed reduction in matrix flow with depth	Does not model Bio, Mech, GeoChem. No DP transfer. ADP modelled by over-simplistic geometry (one channel only)
	SUTRA	McCreanor and Reinhart (2000)	Spatially stochastic waste model	Does not model Bio, Mech, GeoChem. No DP or ADP transfers. No correlation structure to permeability (will end up with AD model asymptotically)
Coupled		White et al. (2004)	Anaerobic biodegradation, mechanical consolidation, Darcian flow	No DP or ADP transfers. Highly parameterised
Coupled		McDougall and Silver (2005)	Deg, mechanical consolidation, Richards' equation for unsaturated flow	No DP or ADP transfers
Coupled		Yildiz et al. (2004)	Phased-construction of new cells. Mechanical consolidation, biodegradation, darcian flow (saturations tracked)	No DP or ADP transfers. Highly parameterised
Risk	SUFI-2	Abbaspour et al. (2004)	GLUE-like refinement of parameter distributions	Relies on Hessian. Determination of priors is difficult.
Risk		Zacharof and Butler (2004a&b)	Bio as well as Stochastic-advective AD model. Testing against data from Brogborough and by Rosqvist and Bendz (1999)	Does not model: Mech, GeoChem. Independent and uncorrelated parameters in Monte Carlo
GeoChem	PHREEQC	Sabbas et al. (2003)	Thermodynamic-equilibrium geochemical model	GeoChem only
GeoChem	Microql	Johnson et al. (1999)	Thermodynamic-Equilibrium geochemical model	GeoChem only

9.3.4 Work needed

The literature review shows that the conceptual model for flow and transport in wastes is still far from certain and indeed may vary from case to case. There is a need to design investigations to

better constrain this uncertainty and more systematically map the possibilities that can arise. There is a need to better understand how justified different simplifications are and to examine the sensitivity of prediction to this uncertainty.

In response to these needs, modelling arguably should use a coherent framework of models as well as simply adding embellishments to particular representations. The framework presented in Chapter 3, was developed in response to this need. New forms of risk models will need to be developed in the future in order to handle *model* uncertainty, since most existing risk-models primarily handle *parameter* uncertainty.

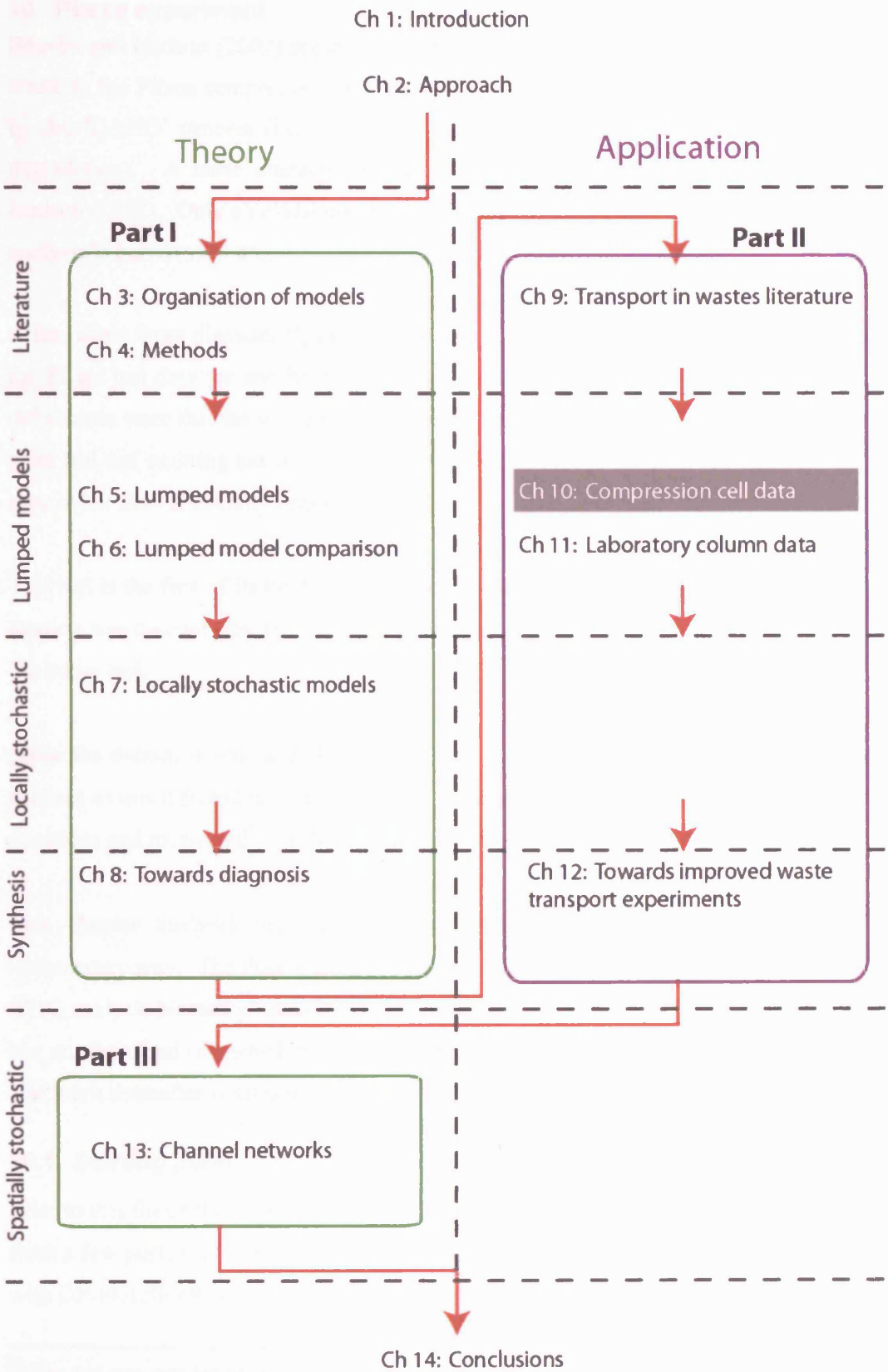
Experiments and models need to break the system down in a reductionist way in order to tease out each process more precisely. The focus here is on conservative transport recognising that further processes need to be incorporated at a later stage.

9.4 *Summary*

Modelling of transport in wastes covers a wide range of possible processes and conceptual models which may be applicable in different cases. For most wastes the pattern of flow is poorly described, but there is often evidence that some preferential flow exists, even if exactly *how* is not known. How to include this preferential flow and how to represent heterogeneity are believed to be of considerable importance to transport modelling. It is useful to find ways in which the system can be simplified. Various questions about this need to be addressed, for example:

- Does the situation when all the processes are operating really need to be modelled? Van der Sloot (2005) suggested that, at least for a starting point, well-degraded wastes are considered.
- Can rate-limiting processes be found which negate the need to measure less important factors?
- Can some processes be considered to be in quasi-equilibrium?

The applicability of the range of models already proposed for wastes needs to be reviewed critically, further models and heterogeneity representations are needed, and basic questions need to be asked of the conceptual model. The practical implications of these questions need addressing, such as to what extent might flux rate or pulsing of flow through waste enhance the flushing.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

10 Pitsea experiment

Beaven and Hudson (2002) presented results from a 146-day tracer test through gassing MSW waste in the Pitsea compression cell (Pitsea landfill, Essex).[62] The waste had been pre-treated by the 'DANO' process (i.e. tumbled in a large diameter drum in order to promote aerobic degradation). A basic characteristation of the waste composition is given in Beaven and Hudson (2002). Only a provisional examination of the data was performed, so a more thorough analysis is performed here.

A few other large diameter 'lysimeter' tests are reported for wastes. A broad comparison with the Pitsea test data for another large diameter cell is made in Rosqvist et al. (2004). The key differences were that the waste used by Rosqvist et al. (2004) was 22-years old, so considerably aged and not emitting extensive amounts of gas and the lysimeter was built into the waste in-situ, rather than artificially packed.[63]

This test is the first of its kind – a controlled tracer test in a large diameter compression cell for gassing waste under positive hydraulic pressure. It also included a period of no flow as part of the tracer test.

Since the dataset is unique and was expensive to collect, a key aim of the analysis here is to squeeze as much from the data as possible. The analysis of the data threw up a large number of questions and motivated[64] much of the theoretical part of the thesis (Part I).

This chapter analyses previously collected data which has already been analysed in a rudimentary way. The data is analysed here in a systematic fashion, by proceeding through the 'BTC analysis hierarchy' introduced in Chapter 4 (Table 4.1). Firstly the basic conditions in the test are described (for which the work of others is described in sections 10.1, 10.2.1 and 10.2.2). The work thereafter is all new.

10.1 *Starting point*

Prior to this thesis the Pitsea tracer test data had not been systematically analysed, but the BTCs from a few ports (out of the 13 outlets) had been randomly selected and provisionally modelled with DP-PULSE (Beaven et al., 2003).

---

[62] The test was conducted by Richard Beaven and Andrew Hudson from Southampton University, with the assistance of Tim Atkinson from UCL. The tracer test and modifications to the cell to allow it were designed by Tim Atkinson and John Barker at UCL in conjunction with Hudson and Beaven

[63] Rosqvist et al. (2004) used cumulative flow to compare the Pitsea data (for Ports O1 and I1) to the data from Rosqvist and Bendz (1999). On cumulative flow axes the BTC for port O1 looked considerably retarded in comparison to port I1, which is potentially misleading (because as will be shown it was not retarded when plotted on a time axis).

[64] The thesis is not presented in chronological order since the transport models and methods are best introduced as part of the theoretical development.

Before the tracer test, the waste hydraulics were characterised by Hudson et al. (2001). They noticed that gas accumulates in the waste and displaces liquid reducing permeability, K (as would be expected for multiphase flow). Increased fluid pore pressure reduces the porosity occupied by gas and is assumed to compresses the gas and enhance dissolution. Increasing the pressure increases gas production rates, too. Once the gas has built up, increasing the water flux rate only had a relatively weak effect on reducing the gas content. This apparently indicates the presence of stable gas-pockets which are difficult to dislodge by liquid pressure. It is possible that in the event of steady flow there may also be an element of dynamic equilibrium, whereby gas removal balances production.

Hudson et al. (1999) showed that there was a possibility that the hydraulic properties of landfills would not vary greatly with depth as there was relatively little difference between the expected properties at the top and bottom of the saturated zone of a landfill. However the paper by Hudson et al. (2001) modified this finding, as there seemed to be a very significant effect of pore pressure on gas accumulation and therefore permeability.

The key observations of this tracer test by Beaven et al. (2003) were: (i) a rapid first arrival of tracer, taken to indicate substantial preferential flow; (ii) a reduction in tracer concentration measured at the outlet of the cell following a 19-day pause in flow which was interpreted as strong evidence of DP transfer process taking place; (iii) a 'long' tail of the breakthrough after the full 146 days of the experiment, and recovery of only 64% of the injected tracer, indicated that some of the tracer had somehow been retained. The DP-PULSE model (Barker et al., 2000) achieved an encouragingly good visual fit to the breakthrough data for three of the outlet ports, using realistic parameters and simple DP geometries (slab and sphere). A constant average flow rate was assumed, except for two periods of zero flow. No account was taken of the inlet or outlet transfer functions of the experimental apparatus.

10.2 *Description of experiment*

10.2.1 Pitsea cell

The Pitsea cell is a 2 m high compression cylinder which is designed for hydraulic measurements of wastes under different compression and pore pressures. The original design is described in Beaven (2000). The cell can be rotated so that different wastes can be placed loosely inside and then manually compacted before the hydraulic ram on the upper platen is used to establish the required working compressive stress. There is a system of tanks and pipework to allow a flow of leachate through the cell. The cell is now housed inside a specially-built superstructure (Figure 10.1). The hydraulic configuration is shown in Figure 10.2.

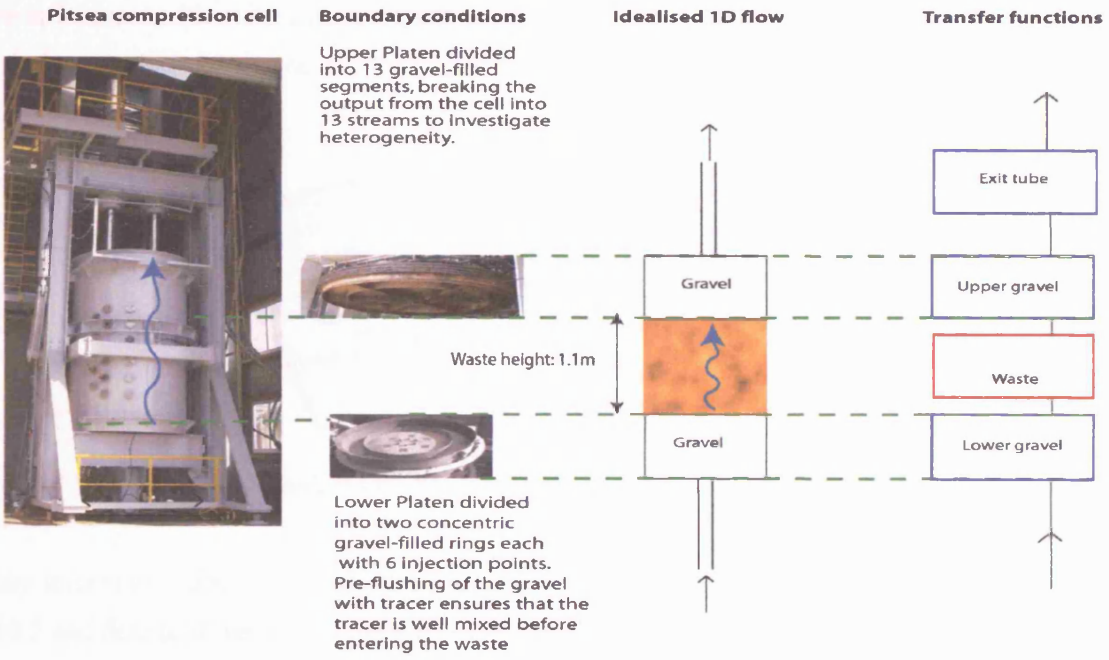


Figure 10.1: Pitsea cell with platens detailed and the idealised 1D representation.

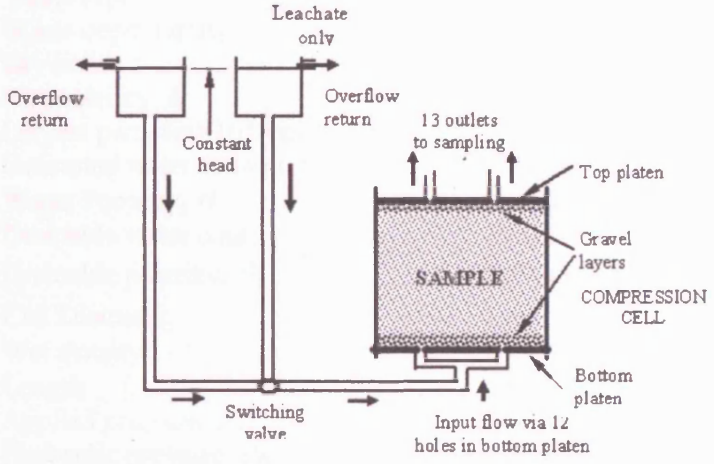


Figure 10.2: Hydraulic configuration for tracer test (from Beaven and Hudson, 2002).

10.2.2 Tracer test conditions

The tracer test is described in detail in Beaven and Hudson (2002). The test was performed to simulate conditions near the base of a landfill with the cell compressed to 228 kPa (equivalent depth of 25 m, assuming wet density of 1.13×10^3 kg /m^3). The hydrostatic pore pressure at the base of the one metre high waste column was 70 kPa, which is equivalent to 7 m below the 'leachate table' (see Figure 10.3). Upward leachate flow under a constant head was established, and then lithium bromide (LiBr) tracer was injected in for approximately four days. There was then a pause in the injection and the cell had zero flow through it for approximately 20 days. This was an unintentional pause in the flow caused by a site power cut. The injection was resumed for approximately one day and then followed by a long period of flushing with clean

water. There was also a second pause in flow for 31 days (starting at 95 days). This was for operational reasons (the site had to be left unmanned for this period). Overall, the BTC was measured over a 146-day period.

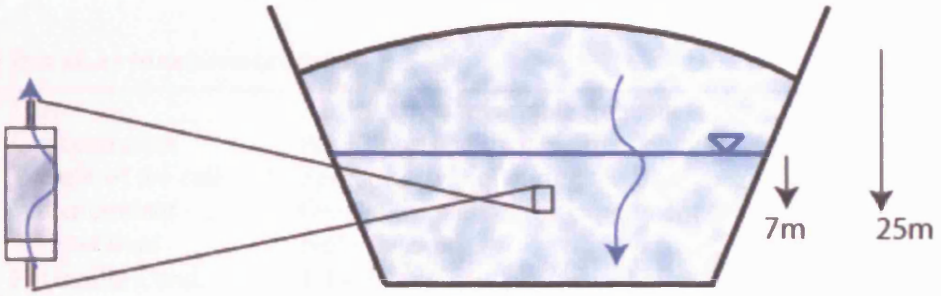


Figure 10.3: Field condition simulated by Pitsea tracer test.

The key information about the experiment are given in Box 10.1, the tracer injection details in Box 10.2 and details of the measurements made in Box 10.3

Box 10.1: Key information for Pitsea cell

Waste type	Gassing DANO[65]
Waste depth (start)	1.10 m
Dry solids	2788 kg
Permeability, K	$10^{-7}<K<10^{-6}$ m/s
Largest particles>165mm	9.1%
Estimated water content	1112 L
Waste Porosity, θ	32.3% (v/v)
Drainable water content	70 L
Drainable porosity, θ_d	2.0% (v/v)
Cell Diameter,	2 m
Wet density	1.13×10^3 kg/m^3
Length , L	1.1 m
Applied pressure, P_{tot}	228 kPa
Hydraulic pressure, P_W	70 kPa
Tracer	Li
Background Conc^n	0.18mg/L Li (or C/C_0=0.007)
Injection Conc^n (C_0)	26.2 mg/L Li
'Target' flow rate	5 L/h
Liquid vol. of upper gravel	70 L
Liquid vol. of lower gravel	120 L

Box 10.2: Injection details

First injection vol.	300 L leachate plus Li^+ tracer
Length of first injection	3.93 days
First pause	19.95 days
Second injection vol.	315 L leachate plus Li^+ tracer
Length of second injection	0.96 days

[65] DANO waste is MSW which has been subjected to the 'DANO' process (tumbling in a large drum to encourage aerobic degradation)

Total liquid injected	5465.9 L
Total flowing	96.0 days
Total paused	50.9 days
Total Time Injecting	4.9 days
Total Time Washing	91.1 days
Total vol. washed	4850.9 L

Box 10.3 : Measurement details

Flow	All 13 ports, at variable frequencies
Concentration	All 13 ports, at variable frequency (not during flow pause)
Weight of the cell	Sporadically, not during the flow pause, some drift
Water content	Only at the start of the experiment
Temperature	Not measured
Hydraulic Cond.	1.1×10^{-6} m/s low gas, taken before the test
	1.1×10^{-7} m/s high gas, taken before the test
Gas production	Not measured

Waste size analysis and categorisation showed that 91.9% of the waste by weight was smaller[66] than 165 mm, with the modal class 165-80 mm holding 39% of the waste. There was therefore 9% by weight of particles which were larger than 165 mm and these are potentially able to generate a significantly heterogeneous response at the scale of the experiment.

The Pitsea cell was specially modified for the tracer test. Firstly, to make a clean 'top-hat' pulse of tracer injection appear evenly along the base of the cell, the lower platen was divided into two concentric segments, in order to allow a flushing procedure to ensure a uniform application of tracer to the base of the waste. Secondly, the top platen was divided into segments.

The efficacy of the procedure to spread tracer through the gravel was tested by a NaCl tracer test and by modelling. Barker (personal communication) used a boundary integral method with the inlet ports modelled as point sources and sinks. Plots of flow streamlines (Figure 10.4) and histograms of flow times were produced. This showed a bimodal distribution of times, corresponding to the long and short paths between ports. There were always some particularly slow paths corresponding to stagnation points and it was concluded that no static pattern of pumping would flush the whole system. A procedure of rotating the pairs of ports was therefore devised to ensure that no stagnant patches remained.

[66] This was presumably subject to what extent particles can squash up to pass through sieves and screen and then spread out once under pressure.

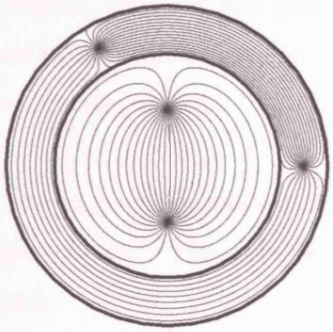


Figure 10.4: Static gravel flushing pattern (from Barker, unpublished), with two pairs of ports 'activated'.

For the main test, the lithium bromide tracer that was selected had been previously successfully demonstrated in MSW wastes (Blakey et al., 1998; Rosqvist and Bendz, 1999). It was intended that both Br^- and Li^+ ions would be measured, with the possible benefit that the different diffusion coefficients of the two species would reveal information about any diffusive exchanges within the waste. Suspected flow and transport heterogeneity in the waste was to be investigated by the division of the top platen into segments, along with an outer peripheral segment to provide a check as to whether the walls of the cell affect the flow locally (either by direct flow along the wall or from smearing of the waste along the wall during packing and compression).

The top platen of the cell is divided into a series of segments, providing a measurement of the spatial variability of transport (Figure 10.5).

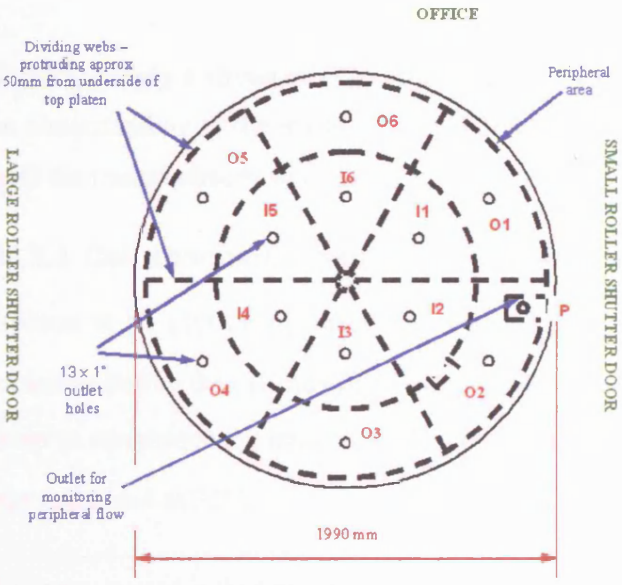


Figure 10.5: Top platen (from Andrew Hudson).

The segments are 116 mm deep and filled with gravel, to allow a uniform flow to or from the surface of the waste. Fluid exits through 10 mm outlet tubes connected to a single port in the centre of each segment.

The flow and concentration measurements were taken by collecting leachate from each tube into storage vessels. The volume and concentration in each vessel were recorded at variable intervals.

There were difficulties with the Br^- measurements. Specific ion electrode measurements taken on site were discarded as unreliable and there were difficulties with using High Pressure Liquid Chromatography (HPLC) following acid digestion of the sample.[67] Since Stegemann et al. (2006) showed that there is significant non-linear Br^- sorption to waste, Br^- may anyway have proved a relatively difficult species to model. This is an interesting finding, as Br^- is assumed in some media to be conservative (Flury and Flühler, 1995). Ward et al. (1998) found that out of the available 'solute-type' tracers for aquifers and soils, the halogens are most suitable, of which Br^- is often preferred (as it often has lower and stabler background levels). So, Li^+ was used as the only tracer. Li^+ was analysed by Atomic Absorption Spectrophotometer.[68]

Stegemann et al. (2006) showed that there was no lithium sorption to an aerobically degraded waste. Deborah Reinhart (in her 1989 PhD thesis), as cited in Öman and Rosqvist (1999) reportedly found no sorption of Li^+ on fresh and acid phase MSW, and a slight affinity in stable waste. This is interesting since Ward et al. (1998) noted that for soil and groundwater use Li^+ can be prone to ion exchange and adsorption and is less suitable than an anionic tracer. The pH is likely to be an important control on sorption.

There is clearly a strong need to provide a basic appraisal of tracers for use in wastes, linked to an understanding of the reactions with different waste components. Until then, the best is made with the tracers already discussed.

10.2.3 Cell conditions: gas and water contents

Hudson et al. (2001) measured drainable porosity, θ_d, for different compressions and pore pressures before the cell was filled with leachate (from the base upwards to avoid air pockets) in order to commence the tracer test. The θ_d at the start was measured as 2% and θ (water filled) was estimated as 32%.

[67] Dionex, located in the Civil and Environmental Engineering Laboratory, University of Southampton.
[68] Spectr AA-200, Varian, located in the Civil and Environmental Engineering Laboratory, University of Southampton.

Cell weight was recorded throughout the tracer test, allowing calculated water-filled porosity to be plotted against time (Figure 10.6).[69] The saturation level (water-filled porosity) is simply calculated as the volume occupied by water divided by the total volume: $(W - W_D)/(V\rho_L)$, where W is the mass, W_D is the dry mass (obtained by drying the waste in an oven at 105 ^{o}C), V is the cell volume and ρ_L is the density of the leachate (assumed to be the same as water).

Beaven et al. (2003) used this θ_d (2%) as the starting value of mobile porosity, θ_m, for calibrating the DP-PULSE model against the breakthrough data. Initial calibration of the model then yielded θ_m above this (between 4% and 10%). The uncertainty in the weight measurements (and hence the porosities) means that it is not possible to clearly show whether this increase is real (due to a different gas content or a different pattern of gas distribution), or whether this is just an artefact of fitting the model. It is not certain whether θ_d can be reasonably used to bound θ_m. It has been suggested that it might be a lower bound for θ_m (William Powrie, personal communication). This is on the basis that some of the pathways may not be drained fully due to perched liquid remaining on top of low permeability objects, and films of liquid remaining attached. Conversely it could be argued that drainable porosity might be an upper bound, since what is drained could include stores of liquid that are attached to the main pathway(s) but do not normally participate in flow. In conclusion, without more direct measurements of the system it is difficult to use the drainable porosity measurement for anything other than a sensible starting value for θ_m. Since drainable porosity is relatively easy to measure, it perhaps could be used as one of a number of indicators that can be compared at the beginning and end of a tracer test to check whether the waste hydraulics changed.

The water content was subject to a number of disturbances, principally two major disturbances corresponding to periods where the flow was stopped.[70] In the long period of wash between the first and second pauses, the water-filled porosity settles to a constant value of approximately 29%. In the period where the flow has been re-established following the first pause, there is an approximately 30 day period where the porosity drops from 32% to the stable value of 29%. This is a small absolute drop, but in relative terms it is still a 9.4% fall in water-filled porosity,

[69] Note that there is some concern over drift in the instrument, so the absolute values cannot be trusted (large swings are however useful indications of change).

[70] The first stoppage resulted from a major site power cut. This inadvertently introduced an intermittent flow pattern, which was interesting in its own right. The downside was that the hydraulic rams no longer delivered their set pressure and following the loss of power there was a sudden loss of compression, sufficient to bend three acros that had been placed as additional support for the top platen (Andrew Hudson, personal communication). The second stoppage was planned and compression was maintained.

which would potentially cause a proportionate increase in the flow velocity and in the porosity ratio, σ.

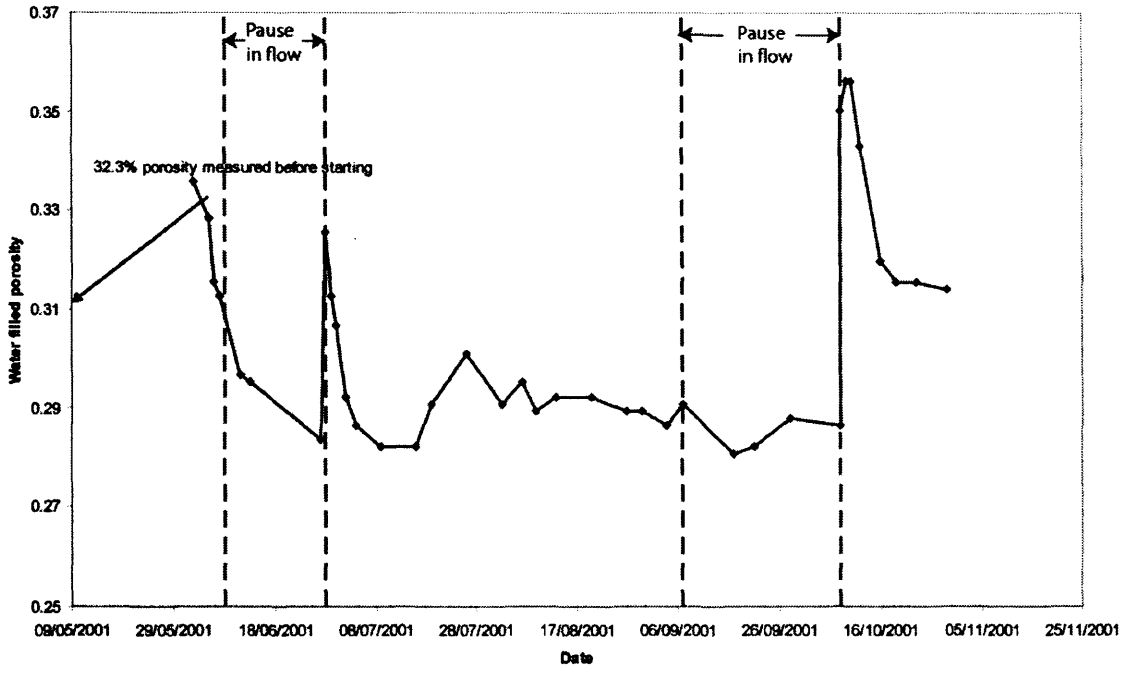


Figure 10.6: Water-filled porosity in cell during test

On the second pause the porosity increased to 36% and then dropped again, but apparently towards the higher stable value of around 31%. The load-cells are notoriously problematic (Andrew Hudson, personal communication); they are reported to 'drift' over time, possibly in response to incorrect temperature corrections. Caution therefore needs to be applied to the absolute values of water-filled porosity reported based on the load cell, though the existence and approximate magnitudes of sudden changes are expected to be more reliable.

It is possible that this porosity is not shared equally by the different stream-tubes. This may explain why there is a noticeably high variance between the different ports during this 30 day period, which tends to a low fairly constant value at the end of this period.

10.2.4 Cell conditions: flow rates

Figure 10.7 shows the average flux in each port, q, and duration of flow. Although the ports were intended to be equal areas, due to various practical issues they actually range from 5.7% (for O3) of the total area to 11.2% (for P). There was a great range in flux (the ratio of maximum to minimum flux was 28.4, the SD of fluxes was 14.6 L/m^2/d). The flow appears to be highly heterogeneous at this measurement scale. What is more difficult to tell is what causes

this; whether the waste is internally highly variable, or whether the variation in flow rate is due to interaction with the outlets. For example, it might be the case that the flow rate variability is caused mainly by choking by gas bubbles at the outlet ports, or is controlled by local geometry of the waste at each outlet (Figure 10.8). Although there are 155 mm barriers to separate the outlet ports, it might still be possible that there is some lateral flow just below each segment, perhaps caused by gas blockage.

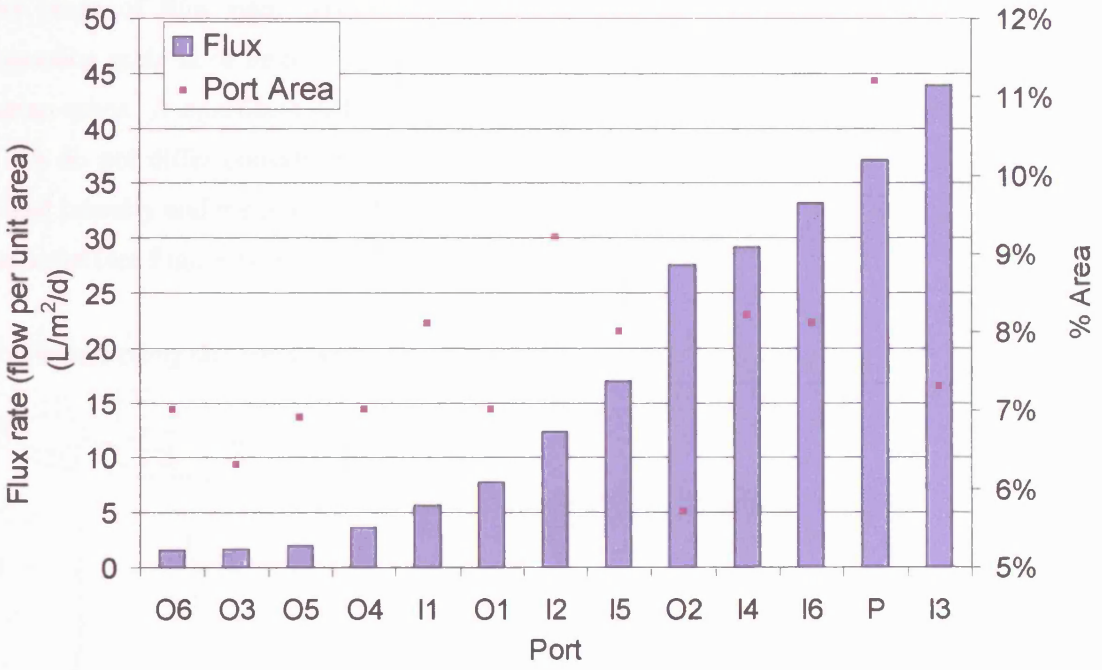


Figure 10.7: Flux rate averaged over whole flow period (bars) and % area of upper platen (squares). See Figure 10.5 for location of Ports.

Beaven et al. (2003) observed that the greatest individual *flow* was through the peripheral port, P (here it is shown to be the second largest *flux*). This apparently demonstrates a degree of enhancement of flow due to the presence of the wall. However, cumulatively more flow passed through the inner ports (2586 L) than through the outer ports and peripheral area (1406 L), so this effect appears not to dominate. Beaven et al. (2003) even suggested that P may scavenge flow from the adjacent O ports, which in general have much lower flux than the I ports. This seems reasonable conjecture, but needs to be confirmed with direct measurement.

When the experiment was refilled at a later date with a shredded MSW waste, P again stood out giving the lowest flow for an 87 kPa low pore water test without gas but then dominating the flow (up to 94%) for the same conditions but with gas accumulated. In wide experience of many flow tests of different wastes under different conditions, Hudson (personal communication) finds that P is often observed to have significantly higher flow than the other ports.

It is interesting to note that there is an imperfect trend for the small area ports to have the lowest flux rate and the largest area ports to have the highest flux rate. Simple calculations of head loss in the different pipes shows that this is not due to variable head across each. Perhaps the larger ports effectively sample a larger number of possible discrete outlet points or lines on the outlet surface, so have a greater probability of draining a major pathway.

The range of flow rates between the ports might be evidence that there is a mechanical dispersion scale at or below the size of the ports, due to different flow velocities in different stream-tubes. A significant difference in the BTCs would be one way of confirming this. If the BTCs do not differ considerably, then a second explanation could be that the system is well mixed laterally and the major difference between the ports is just down to local flow control at the outlet (see Figure 10.8)

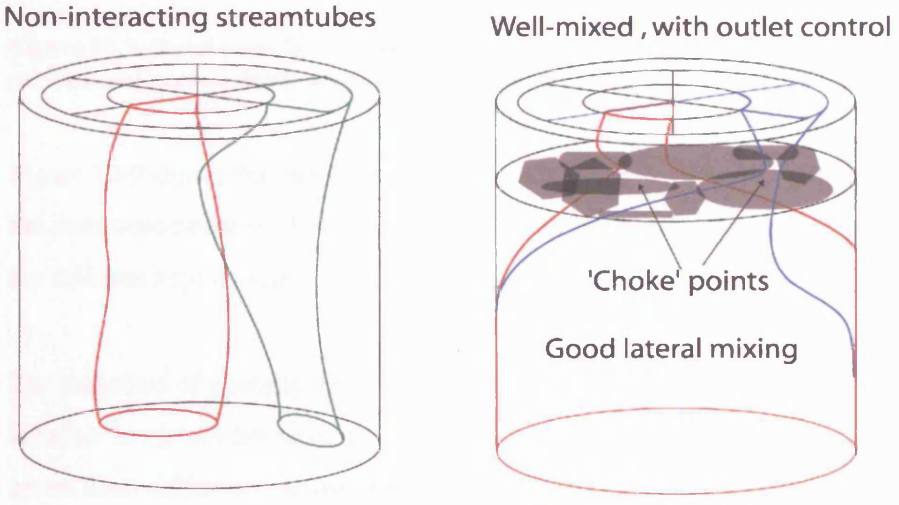


Figure 10.8: Different explanations for outlet flow variability.

More local averages than the flux found in Figure 10.7 can also be found from the cumulative flow for smaller and smaller time periods, down to the time between measurements (Δt). Because of the constraint of model run times, the effect of choice of varying Δt from 0.1 days to 146 days on the model outputs has not been systematically investigated. The approach taken here was to select stress periods of constant flow, separated by sudden changes in gradients detectible on the cumulative flow graph (Figure 10.9).

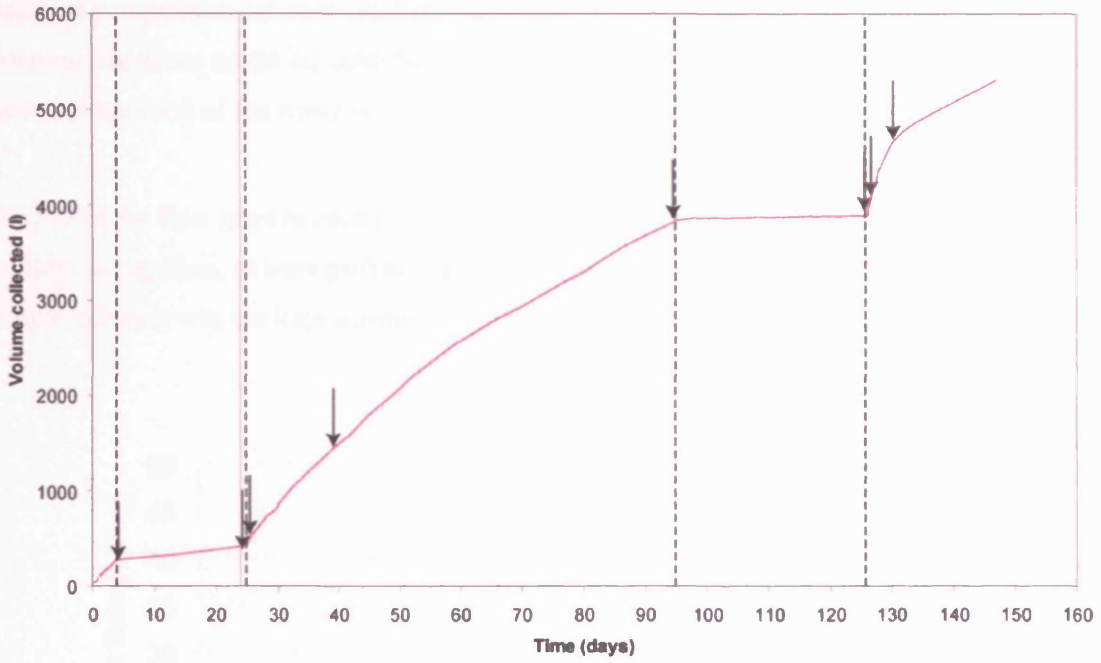


Figure 10.9: Total cumulative flow for the Pitsea tracer experiment (dotted lines mark pause periods and arrows mark assumed constant flow stress periods).

Figure 10.9 shows that the flow rate varied with time, particularly during the few days following the commencement of flow at the beginning of the test and after both pauses. The head across the cell was kept constant, so K (or relative K) must have been varying with time.

The variation of porosity against time visually correlates with the flow rate (however, r^2=0.37 is rather low), at least in terms of the timing of major spikes. This would be expected as for a given head difference across the cell, the higher the gas content, the lower the K to leachate. It would be expected that the variation in porosity determines K, and hence the flow rate for a constant head. However, any possible relationship between θ and K has not been established for this medium.

This concludes the description of the experiment. Hereafter, this chapter provides an analysis of the breakthrough curve (BTC), organised according to the 'BTC hierarchy' of Table 4.1.

10.3 *Level 1 (Transport connection)*

Figure 10.10 shows that Level 1 in the BTC hierarchy was often met, although there may be intermittent periods of no-flow and therefore no-transport.

This effect may be due to internal instabilities, such as those described by Glass et al. (2002) – multiple switching pathways (in unsaturated fractured rock) with fluctuations of the timescale of minutes to hours. Glass et al. (2002) suggested several mechanisms; the presence of a finely

balanced competition between capillary, gravitational, viscous and inertial forces combined with clogging bio-films could all contribute to an unstable flow. Path-switching may mean that a greater proportion of the waste is flushed than would have been with a more static flow.

The SD of the flow rates in each port is correlated (r^2=0.76) to the flow rate. The noise model for flow is therefore, at least partially, a relative one. The fact that the sampling frequency (and sample volume) was not kept constant further complicates the picture.

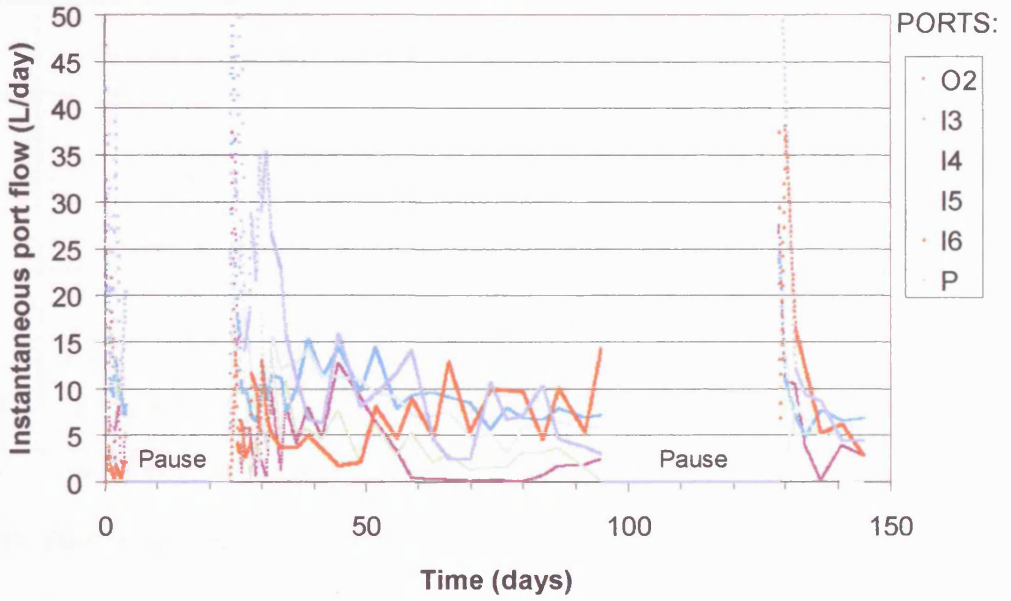


Figure 10.10: Flow rates in selected ports

10.4 *Level 2 (Key features)*

10.4.1 Level 2 (Key features: arrival time)

Looking at the BTC for early time at each port, reveals the time which the tracer was first detectable above background at each port. This observation allows a lower-bound estimate[71] of θ_m. Assuming that the flow to each port passes through a volume fraction of the cylinder equal to the fraction of the platen area occupied by the port, this gives a range in θ_m from 0.016 (I3) to 0.0009 (I5). This is calculated by $\theta_m \geq \frac{t_1 Q}{AL}$ and is a lower bound because $t_1 < t_a$.

[71] It is difficult to relate the measurement of 'first detectable' tracer to any other parameter values, as it is a value which is dependent upon the detection limits, background concentration and the sample frequency.

Figure 10.11 shows the arrival times in order of increasing central time[72]. There is a weak negative correlation (-0.53) with the average flux rate up to the first pause in each port. Whether a time is well constrained or not depends on the sampling frequency for a particular port. Ports O6 and O3 were relatively infrequently sampled.

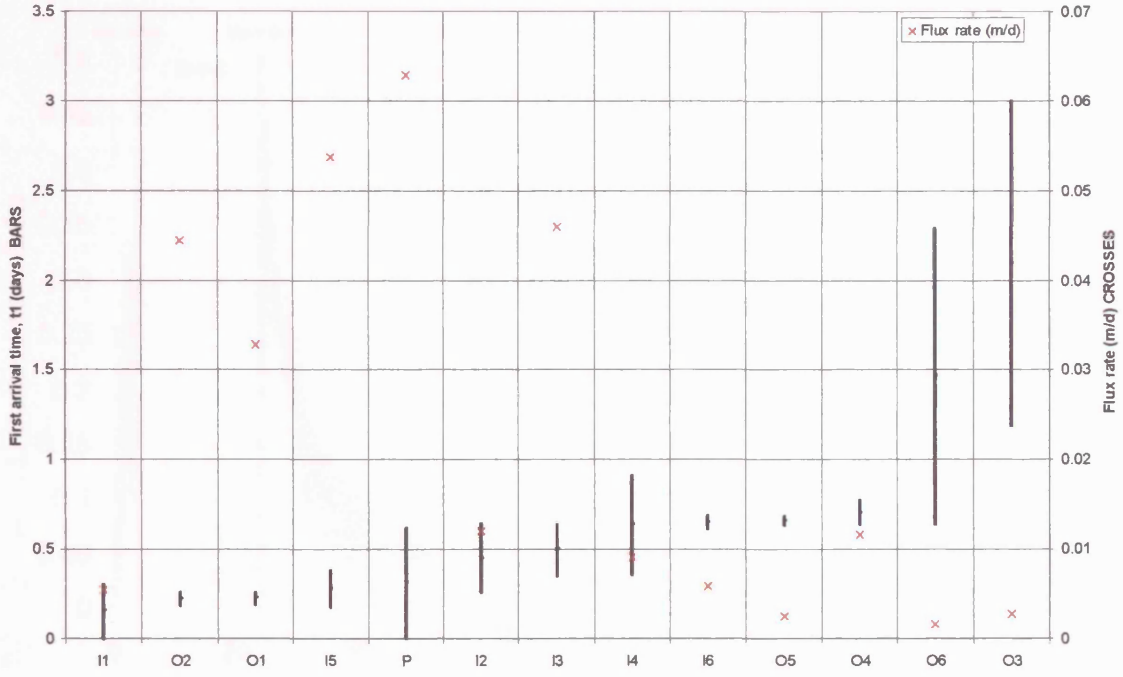


Figure 10.11: First-arrival time and flux rates in each port.

Beaven et al. (2003) noted that, "the rapid breakthrough of tracer provides evidence of channelling". Since the AD model works on an assumption of a range of travel times, potentially produced by the existence of many intersecting fractures or channels, the presence of 'early' arrival does not preclude the existence of a highly laterally dispersed AD behaviour. Furthermore it does not prove (or disprove) 'preferential flow' in the sense that flow is dominated by one fast 'channel'. Therefore if the dispersivity is set high enough, it is possible to model 'early' breakthrough with a basic AD model although it is questionable whether dispersivities as high as 10 m or so are physically plausible for a 1 m long column.

10.4.2 Level 2 (Key features: BTC)

Having examined the early-time in its own right this section considers the whole BTC. Box 10.4 provides the key information about the BTC. If the ports had been connected together, by conservation of volume, the blended concentration would be given by:

[72] The central time is calculated as the midpoint between the maximum and minimum time that can be deduced from the two data points either side of a raise in concentration above background.

$$\overline{c}(t) = \frac{\sum_{i=1}^{n} c_i(t) q_i(t)}{\sum_{i=1}^{n} q_i(t)} \qquad (10.1)$$

Figure 10.12 shows the BTC produced by this blending.

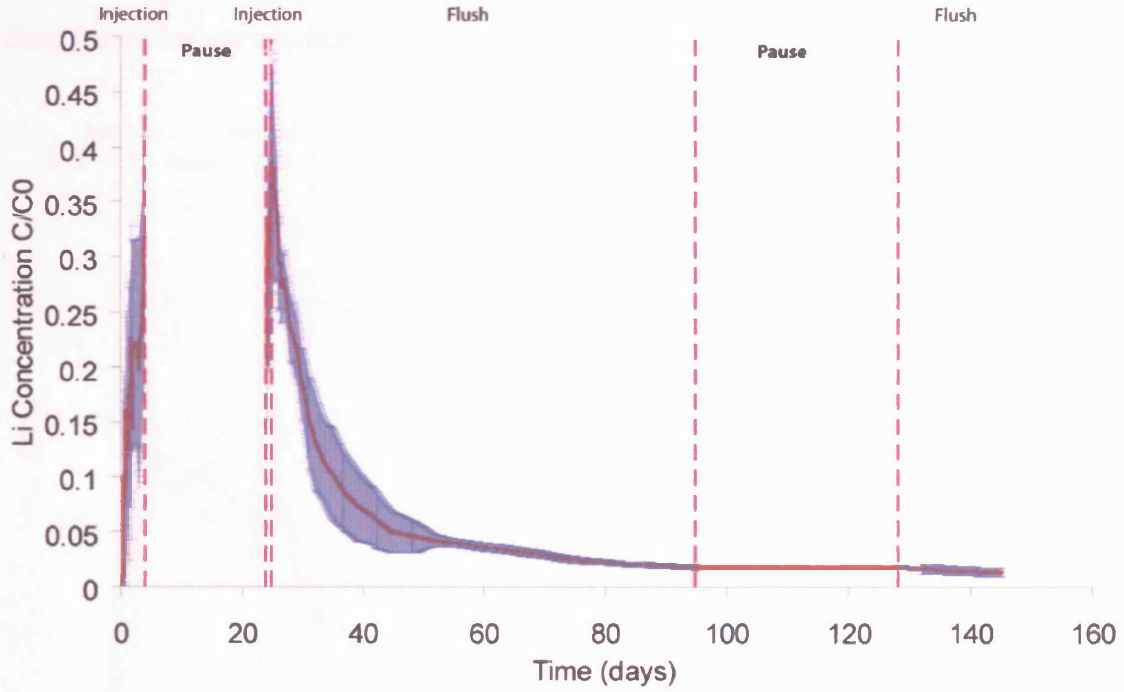


Figure 10.12: Combined BTC and SD of all port data (error bars 1 SD). Notice injection, flush and pause periods.

The input condition was a period of injection, followed by a pause, a further short period of injection and then a period of washing (with a pause at late time). The concentration fell during the first pause, rapidly climbing again on resumption of tracer injection. DP transfer is the most likely explanation of this behaviour and it is difficult to suggest other mechanisms to explain this (for example, the possibility of kinetic sorption has been eliminated by the measurements of Jing-Lin (2005)).

Given that this appears to be a well-behaved BTC it may be amenable to a continuum process model representation.

BOX 10.4: Key data on the BTC.

Start of first pause	3.9 days
Length of first pause	19.9 days
Length of injection after first pause	1 day
Start of second pause	94.9 days
Length of second pause	30.9 days

10.4.3 Level 2 (Key features: BTCs at different ports)

Figure 10.13 and Figure 10.14 show the BTCs plotted for the outer and inner ports respectively. The flow through the outer annulus (labelled 'P' in Figure 10.14) was noticeably flashier, with the steepest rise to the highest peak concentrations as well as the fastest fall in concentration. This reinforces the suggestion based on the flow data that there is possibly sheet flow along the wall, or perhaps through pores space created by vertically aligned deformation, due to smearing along the wall-waste interface.

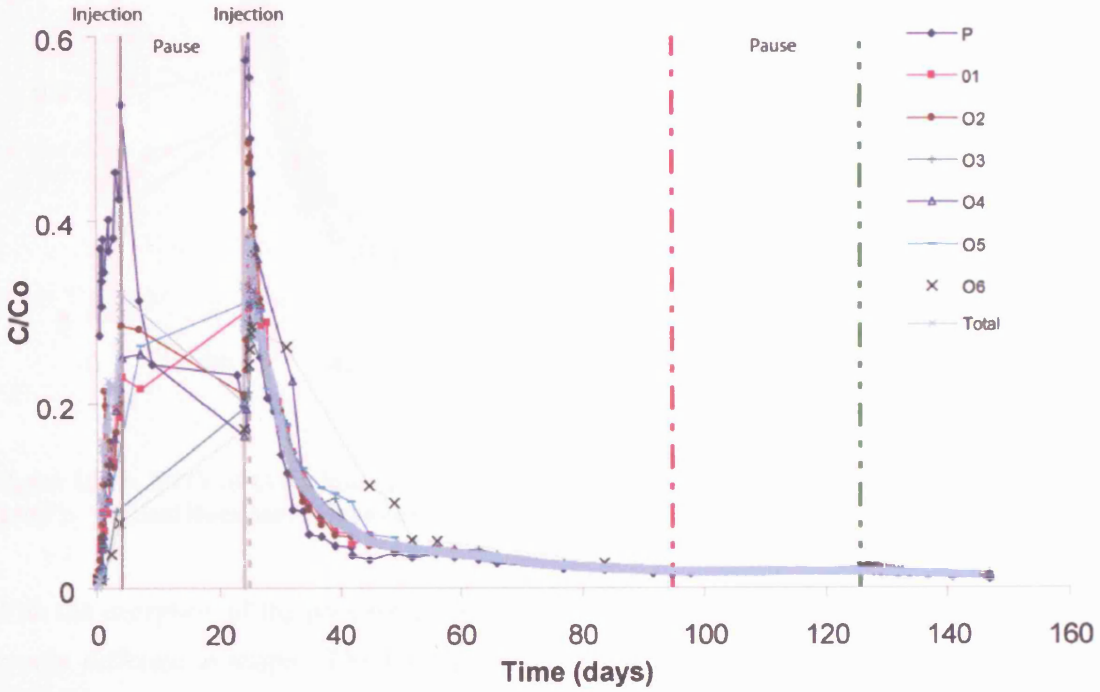


Figure 10.13: BTCs of Li for outer ports (marked O), the peripheral port (marked P) and the synthesised volume-average (marked 'total'). Vertical lines denote injection and pause periods.

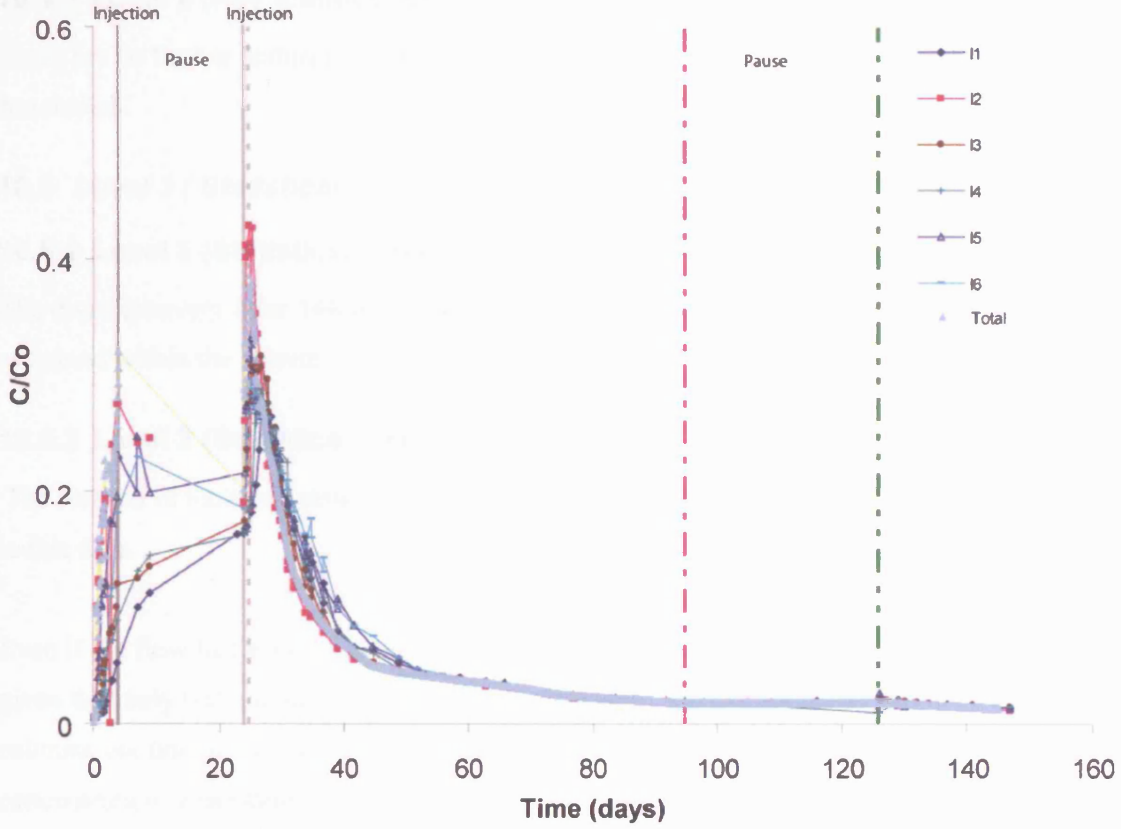


Figure 10.14: BTCs of Li for inner ports (marked O) and the synthesised volume-average (marked 'total'). Vertical lines denote injection and pause periods.

With the exception of the peripheral port and the period following the pause, the BTCs are not greatly different in shape. The assumption that the entire cell can be regarded as a REV, or continuum, especially for periods with steady flow, is not unreasonable. Most mixing could take place at a smaller scale than between the outlet port stream-tubes, or alternatively there is very strong lateral mixing (a notion supported by the expectation that sub-parallel layers of plastic will cause considerable horizontal flow). The waste might be considerably more homogeneous and behave more like a granular medium, than may have been anticipated.

10.4.4 Level 2 (Key features: peak time)

The time after the start of the test that the peak concentration occurred (based on the average concentration for all ports) was at 24.8 days. Unlike a peak time for a BTC that occurs in a test with an uninterrupted constant flow rate, it is difficult to relate this time to meaningful parameters of the system.

10.4.5 Level 2 (Key features: late-time gradients)

The BTC up to the end of the test does not display a concentration range in excess of two orders of magnitude. Such a range in concenetration is arguably insufficient for an adequate 'late-time' analysis of the gradient of the BTC tail.

10.4.6 Level 2 (Key features: other features)

There are no further features of the BTC that are readily apparent. Level 2 is therefore concluded.

10.5 *Level 3 (Statistics)*

10.5.1 Level 3 (Statistics: mass recovered)

The mass recovery after 146 days was 64% . A substantial proportion of tracer therefore still remained within the system.

10.5.2 Level 3 (Statistics: moments)

The method of moments relies on there being an uninterrupted flow, so would not be applicable to this data.

Even if the flow had been continuous it would have been difficult to estimate the moments given that only 64% of the mass had so been recovered. An extrapolation would be needed to estimate the fate of the remaining 36% of tracer. Since moments are highly sensitive to the concentration at late-time, the results would be strongly dependent upon this assumption. This method is best suited to cases where there is a high recovery and when there is a well-established log-lin linear late-time gradient and a good physical justification for its extrapolation. However, the method of moments are later used in Section 10.7.1 in order to assess the effect of the outlet gravel on the BTC.

10.5.3 Level 3 (Statistics: other statistics)

The noisiness of the data, as a function of time, was characterized by the SD of the concentrations at the different ports. Figure 10.12 showed that the SD following the second peak achieves a constant low value at about $t = 60$ days (30 days after the peak). This coincides with a stabilisation in the weight of the waste and also with the reduction in variability of flow between the ports (Figure 10.13). A further factor is that the variability would be expected to be higher at earlier times where unmixed high and low concentrations are likely to exist, whereas later on the concentration is uniformly low. The error in concentration measurement may be partially proportional to concentration (similar to the flow measurement).

10.6 *Level 4 (Modelling: general)*

10.6.1 Level 4 (Modelling: previous calibrations)

Beaven et al. (2003) showed that a DP model fitted the breakthrough curve. They achieved a good visual fit to data from the individual ports 'I2, 'I3' and 'P', assuming a constant flow rate

based on the average for the entire cell, a single pause and no allowance for the transfer function of the experimental apparatus. It is worth noting the following:

- For I3, advection time, t_a, was between 0.56 and 1.40 days, fracture time, t_{cf}, between 0.1 and 0.5 days.
- Block time, t_{cb}, was between 90 and 112.5 days.
- Using $b = \sqrt{t_{cb} D_a}$, where the apparent diffusion coefficient, D_a, is approx 3×10^{-10} m^2/s, predicts block size, b, to be between 4.8 and 5.4 cm. This seems reasonably close to typical waste 'particle' sizes.
- The assumption of constant flow is highly questionable, as can be seen from Figure 10.9. If more realistic variable flow rates are used for the same calibrations, very poor fits result.
- Taking the best-looking visual fit to I2, I3 and P, the goodness-of-fit in comparison to the total concentration curve (by interpolating to the nearest points) gives r^2=0.87 (sphere) and 0.86 (slab).

Beaven et al. (2003) suggested in a preliminary way that DP models might be appropriate. Whether other representations were equally or more valid was not investigated. A systematic calibration was needed to give a more reliable assessment of parameter values. The complete range of ports and their combined effect also needed to be considered beyond this random selection of three. This need for a more thorough analysis following this 'pilot' formed the starting point for the work here.

More specific questions emerged. What was the effect of the transfer function of the rig, is it important and can it be bounded? Did the variable flow rate affect parameter values or process identification significantly? How best should heterogeneity be characterised?

10.6.2 Level 4 (Modelling: assumptions)

The assumptions that were needed for the previous fits and in order to allow further modelling now need to be stated.

The process behaviour of the system is not assumed here to be time-variant (i.e. the system is assumed to be 'stationary'). This is a relatively questionable assumption, because the system is continuously degrading, settling and producing gas and liquid leachates. The condition of the system over the entire experiment can be tested by the following questions:

1. Was there measurable degradation?

2. Was there settlement?
3. Did the total pressure change?
4. Did the flow rate change systematically?
5. Did the weight of the cell change?

The evidence for each needs to be examined.

There was no direct measurement of the degradational status of the waste (only cell weight and waste height data). Gas emission wasn't measured. Pressure was not measured and since the top platen was fixed in place it may have varied. The flow rate was measured.

There is no evidence to support any long-term systematic change at the timescale of the experiment, but the flow rate and cell weight data suggest that the flow system was in a transient state following both periods of no-flow. There is insufficient evidence to demonstrate that there wasn't any change either. Future datasets will need to demonstrate this more clearly.

It is assumed that the gas is in a state of dynamic equilibrium (i.e. able to flow upwards and out of the system through either a connected gas-phase which does not change with time and /or by more intermittent transport in bubbles). Indeed, the upward flow configuration was chosen in order to avoid the complicated situation of gas bubbles floating upwards against a downward moving flow. If the more intermittent release is a dominant mechanism, then it must be assumed that over the timescales of the experiment that the flux rate and effective geometries created by the bubbly structure can be represented by 'averaged' properties.

This hypothesised steady-state gas contributes to the 'pore structure' geometry experienced by the liquid fraction. The gas system is expected to be a complex function of waste type, compression and flow history (Hudson et al., 2001). Whether there is considerable systematic variation within the cell is unclear. Also uncertain is to what extent the cell produces an unrealistic behaviour due to the artificial boundary conditions (a 1 m long section of waste deep within a landfill would have waste above and below, not gravel and tubes delivering or removing leachate).

The lumped transport models used here apply single parameters to the entire column, which can be interpreted in two ways. Firstly, it might be assumed that the column has uniform properties throughout its length. Secondly, any variability (perhaps systematic variability with length, or else random variability) in properties is averaged out and single homogenised values are found by calibration. To know whether the former or latter assumption is more realistic requires both detailed measurements at the sub-column scale and then conceptual models to explain the

effects. There have been no saturation-depth measurements, but Hudson et al. (1999) measured compression at different depths in the Pitsea cell and showed no systematic changes.

10.6.3 Level 4 (Modelling: early time)

Comparison of the early part of the curves (up to the first pause) is done here with the simple analytical models (Figure 10.15). The MIM model fits convincingly (calibrated at θ_{im}=0.35, θ_m=0.02, κ=17.5 days^{-1}, α=0.1 m, t_a=0.5 days), whereas the AD model (calibrated at V=0.01 m/day, α=10 m) and the DP model for a block (calibrated at σ=14.5, t_a=0.25 days, t_{cb}=40.2 days) give very similar fits.

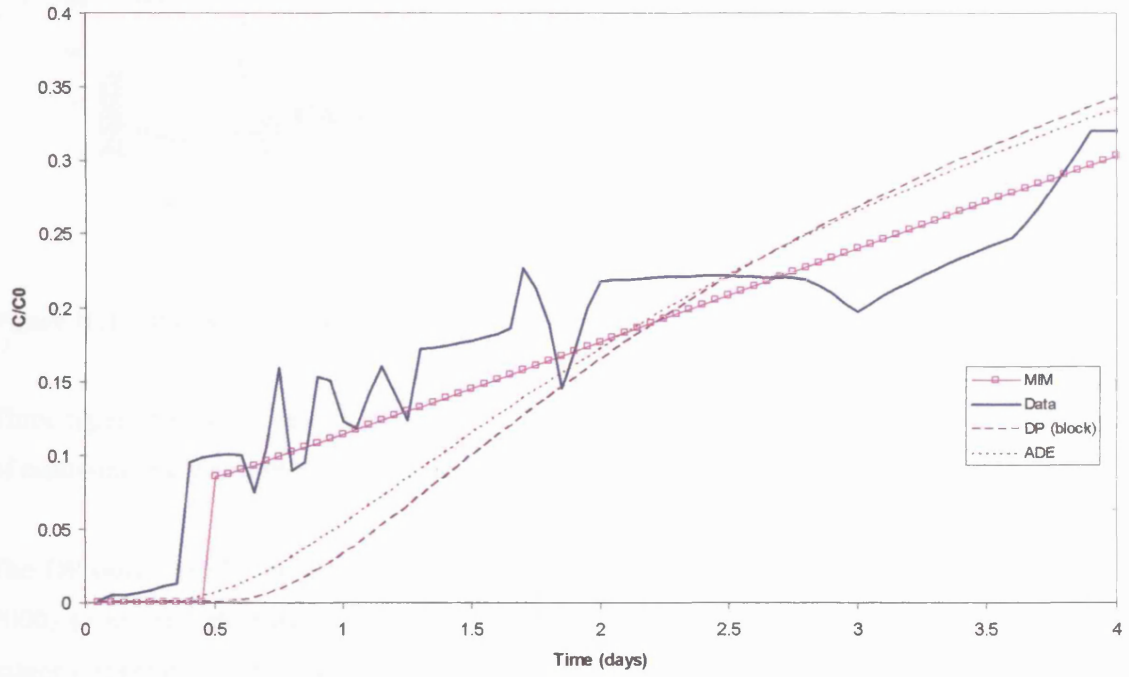


Figure 10.15: Early-time fits to the volume-averaged data for all ports.

The MIM model captures the early jump in concentration. This is not predicted by the DP model, which never predicts a sharp jump or 'shock' in concentration. The early concentration data are very variable between ports and the flow is not very well settled at this point. Without demonstrating repeatability of this data, little confidence should be applied to this data.

A different explanation for the early 'kick' is that a proportion of the solute arrives as film-flow either internally or along the cylinder walls. Very little of this fraction undergoes exchange or dispersion (perhaps it is sandwiched between plastic layers). The later slope is then made up by other stream-tubes that have more exchange-like behaviour.

10.6.4 Coding requirements

Having assumed that transport parameters remain fixed, the only changes to the system are $C_{in}(t)$ and $q(t)$. These are modelled as piecewise-constant inputs (there are two periods of injection and q has been split into seven periods).

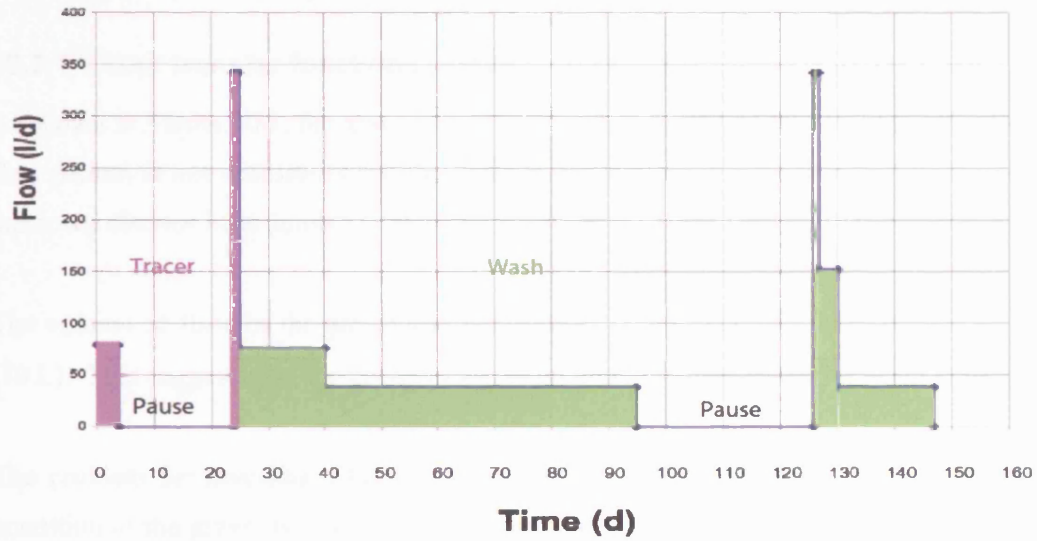


Figure 10.16: Piecewise 'constant' flow determined from gradient changes of cumulative flow.

Three types of models were developed here: DP only, DP-AD and AD only models. The effect of outlet-mixing was added to all three model types.

The DP only model was adapted by Barker from the original DP-PULSE code (Barker et al., 2000) to incorporate variable flow rates. The DP-PULSE code was further adapted by this author (Appendix 1, page 424) to allow mechanical dispersion in the fractures. This hybrid-code is called DPD-PULSE ('D' for mechanical Dispersion). Although the DPD-PULSE code can model AD only, it was decided to build a semi-analytical AD code, which would not have any of the potential numerical errors that can occur for DPD-PULSE under certain circumstances. This also provided a cross-checking mechanism for both codes. This AD code is called ARDVARCS and is detailed in Appendix 1 (page 432).

Automatic calibration was performed using the PEST code (Doherty, 1994). Starting values and ranges for maximum and minimum values that the search routine could try were set for each parameter. Starting values for the DP process were set to the prior values established by the manual fits of Beaven et al. (2003), given in 10.6.1. Wide ranges were allowed for the DP characteristic times (1×10^{-5} to 1×10^{5} days). The starting value for dispersivity, where the AD process was included, was arbitrarily $\alpha = 0.01$. Wide ranges in dispersivity were also allowed

($0.000001 < \alpha < 1000$ m). The final parameter that was varied was V_m. This was also allowed a wide range (1×10^{-5} to 1×10^{5} m/day). Checks were made to ensure that the fits that were obtained did not result in obviously non-physical parameter values. The PEST optimisation parameters were set to the code default values.

10.7 *Level 4 (Modelling: results)*

10.7.1 Other transfer functions (gravel and outlets)

As shown in Figure 10.1, the gravel above the waste is contained laterally by metal plates. The flow pattern is not necessarily simple, perhaps converging on the outlet port from a number of unknown discrete inlet points.

The volume of fluid in the gravel was the same as the volume of drainable fluid in the waste (70 L). This suggests that the gravel might be an important part of the transport system.

The problem for assessing what the process function of the gravel may be, is that the input condition to the gravel is unknown as the junction between the packed waste and the gravel is hidden from view. Even if an experiment to measure the gravel transfer function was run in which a tracer entered the gravel in a uniform plane at the bottom – this would not necessarily represent the situation during the experiment. At best the effect of the gravel can be bounded by applying the most extreme process function that can be envisioned and measuring the effect on the calibrated parameters for the main model.

It is not obvious what flow scenario within the gravel might give rise to an extreme transfer function. A SUTRA model (Voss, 1984) was used to scope the hypothesis that radial convergence might give an extreme behaviour. This was done by assuming a cylindrically shaped outlet port. This allowed use of radial symmetry to simulate a 3D flushing system with a 2D model (Figure 10.17).

The BTCs produced by the pseudo-3D model were very similar to modelling the gravel as a pure mixing cell. A mixing cell is therefore used as the assumed greatest likely perturbation from the system with no gravel (in the absence of any more extreme suggestions).

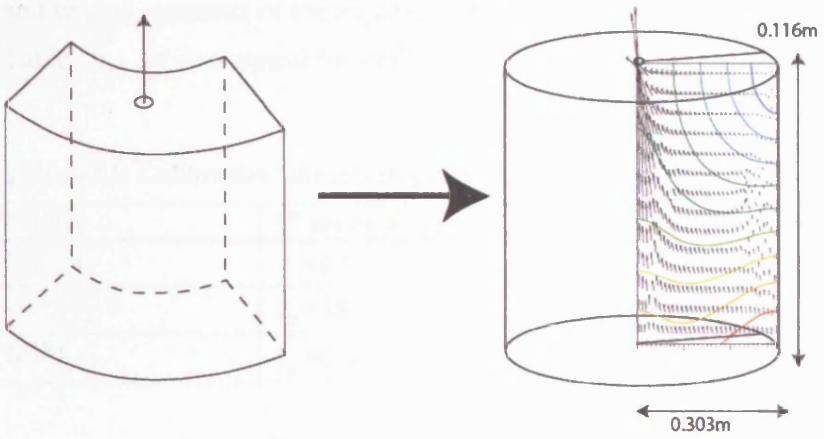


Figure 10.17: SUTRA scoping model for 'worst-case' gravel transfer function. (a) shows flow leaving from an outer port above a gravel 'sector'. (b) shows the SUTRA model, superposed on to a hypothetical radially-symmetrical 'sector'. Flow may enter all along the lower boundary of the model and leave only at the outlet port.

The impulse response of the gravel modelled as a mixing cell is $g(t) = (1/\tau)e^{-t/\tau}$, where $\tau = V_{gravel}/Q$. The first temporal moment of the gravel impulse response is $\int_0^\infty \frac{t}{\tau} e^{-t/\tau} dt = \tau$.

The mean of the impulse response of the combined system (where f(t) is the impulse response of the waste, without the mixing cell) is described by

$$E[C(t)] = E\left[\int_0^t g(\tau) f(t-\tau) d\tau\right] = E[g(t)] + E[f(t)] \qquad (10.2)$$

Note that the increase in the mean time of arrival induced by a mixing cell of the same volume and flow rate is equal to the pure advection time delay through the gravel. The total gravel liquid volume was 70 L and the average total flow through the cell was 55 L/d, making τ approximately 1.3 days. The effect of delaying the tracer by gravel mixing is potentially significant for the calibration of the relatively short-term behaviour, but will become insignificant for the tail. Since the early-time data are of relatively low reliability, this is not much of a constraint.

The spreading of the BTC can also be related using simple probability theory, by:

$$\sigma^2[C(t)] = \sigma^2\left[\int_0^t g(\tau) f(t-\tau) d\tau\right] = \sigma^2[g(t)] + \sigma^2[f(t)] \text{ and } \sigma^2[g(t)] = \tau^2.$$

Thus additional spreading due to mixing is approximately 1.7 days.

Calculation of the complete moments of the impulse response from the BTC cannot be done because there was not complete recovery. Instead, moments can be estimated based upon the fitted process models and calibrated parameters (in the upcoming Section 10.7.2.2). The first

and second moments of the impulse response of the mixing ADE and MIM models are given in Table 10.1 for two typical fits of the AD model and the DP (MIM) model.

Table 10.1: Calibration values (using calibrated values from Table10.2 in Section 10.7.7.2).

Model	1st moment (days)	2nd moment (days^2)
Mixing	τ =1.3	τ^2 =1.7
AD	t_a=18.9	$2t_a t_D$=767.3
MIM	t_a=6.3	$2t_a t_M \sigma$ =479.1

The mixing cell contributes a tiny amount of spread compared to that generated in the waste, thus will make very little difference to the combined output.

The outlet tubes add minimal delay. The outlet pipes were 10 mm ID and approximately 5 m long. This gives a transit time of approximately ten minutes. Given that this time is much less than the key timescales of the test, not to mention the delay time caused by the gravel, it is considered to be negligible.

For the analysis that follows the outlet tubes are ignored and the effect of the gravel is bracketed by assuming a lower bound of no gravel and an upper bound of a mixing cell with a volume equal to the gravel volume.

10.7.2 Model identification

Different models were fitted to the data for the flow-averaged BTC. Specifically the AD model, DP model with various BGFs (including the MIM model) and DP-AD model were fit, all with and without mixing due to the gravel. This analysis was published in Woodman et al. (2005).

The calibration of models to these data was less straightforward than might have been expected. A number of issues arose:

- The codes were slow (i.e. an excess of ten minutes per run for some parameter combinations).
- There was non-uniqueness (different best-fits arising from different starting points).
- There were a similar goodness-of-fits for different models which made bset-fit model identification difficult.
- Weighting had a large effect on the fitted parameters.

The results of calibrating these models are discussed in the following sections on a process by process basis. Figure 10.18 gives the BTCs of the weighted best-fits. Note that non-uniquenss

within each process mean that it is difficult to make meaningful statements about the differences in the BTCs.

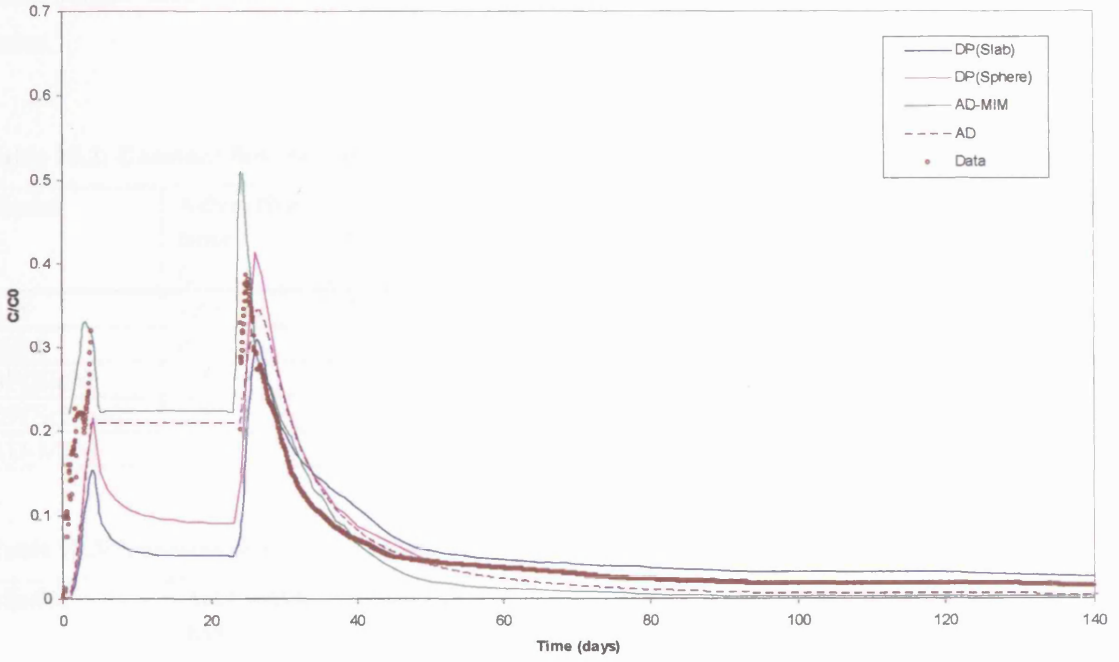


Figure 10.18: Best-fit BTCs for DP(Slab), DP(Sphere), AD-MIM and AD models with constant flow rate. Best fit-parameters are given in Table 10.2.

10.7.2.1 AD model

The AD model failed to simulate the fall in concentration during the pause in flow. However, apart from this it can be fitted reasonably to the data (a worse fit than for the DP-models, but nonetheless it appears to be reasonable visually).

The fit for variable flow rate is better. The best fit occurs for θ>1, which is clearly impossible. Setting an upper bound for θ of 0.323 (i.e. making maximum θ equal to the total measured porosity (Box 10.1)) gives for the SD-weighted variable flow-rate calibration α / L=1.1. As this is high, this suggests that another process is causing at least a proportion of the dispersion, as this is a larger α / L than would be expected (see Section 3.2.1.1). The DP process is a likely further process that contributes to dispersion.

10.7.2.2 DP models

Tables 10.2 and 10.3 give the details of the calibration and Figure 10.19 shows the range in predicted block times. The range in 95% CIs for all the models is from 5.3 to 44.5 days. This shows that at the duration of the experiment (145 days) diffusive disequilibrium is important

and must be taken into account. However, at the timescale of field flushing (order of years) it could be concluded from these times that equilibrium would be established and that DP models would not be required. Although, it is extremely likely that larger heterogeneity scales as well as the existence of lower-saturation towards the top of the landfill would combine to create much longer block times. The variable flow rate causes a reduction in the estimated block times.

Table 10.2: Constant flow rate assumed (key times in days).

Model	Advective time t_a	Dispersive time t_D	Characteristic block time (t_{cb} or $t_{IM}=\theta_{im}/\kappa$)	Characteristic fracture time (t_{cf} or $t_M=\theta_m/\kappa$)	Weighted r^2
AD	18.9	20.3	-	-	0.65
MIM	0.7	-	11.7	0.4	0.82
DP(Slab)	1.1	-	44.5	0.26	0.90
DP(Sphere)	1.1	-	42.8	0.12	0.89
AD-MIM	2.2	0.104	11.8	1.0	0.89

Table 10.3: Variable flow rate assumed: key times (in days).

Model	Advective time t_a	Dispersive time t_D	Characteristic block time (t_{cb} or θ_{im}/κ)	Characteristic fracture time (t_{cf} or θ_m/κ)	Weighted r^2
AD	18.9	20.6	-	-	0.92
MIM	0.9	-	6.3	0.18	0.95
DP(Slab)	1.1	-	20.8	0.039	0.96
DP(Sphere)	1.1	-	10.3	0.038	0.95
AD-MIM	1.1	0.08	5.3	0.25	0.95

The jump in concentration following the pause in flow showed that further processes were occurring in addition to or instead of AD. This could have been expected to have given an increased ability to distinguish between AD and DP processes. Calibration of data against the MIM model only up to the point following the first pause, yielded a block time of 4.7 days, about 75% of the time estimated using the whole curve. The early data are dominated by the effect of the pause, whereas the late time data characterises the longer-term behaviour of the system. Therefore mechanical dispersion probably makes a minor, but not insignificant, contribution to the apparent block time. This is only a semi-quantitative observation, and is somewhat limited by the fact that the early-time data was subject to the greatest uncertainty. Indeed, the concentration immediately after the pause was especially poorly resolved as it was climbing so steeply that a small change in time of sample collection made a big difference to the concentration recorded. The fact that porosity and flow-rate did not remain constant following the pause, further eroded the confidence that could be attributed to the data at this key time.

Figure 10.19 also shows the range in predicted fracture times. These varied over almost two orders of magnitude, from about a day to half an hour. The effect of modelling variable flow-rate was again to reduce the times.

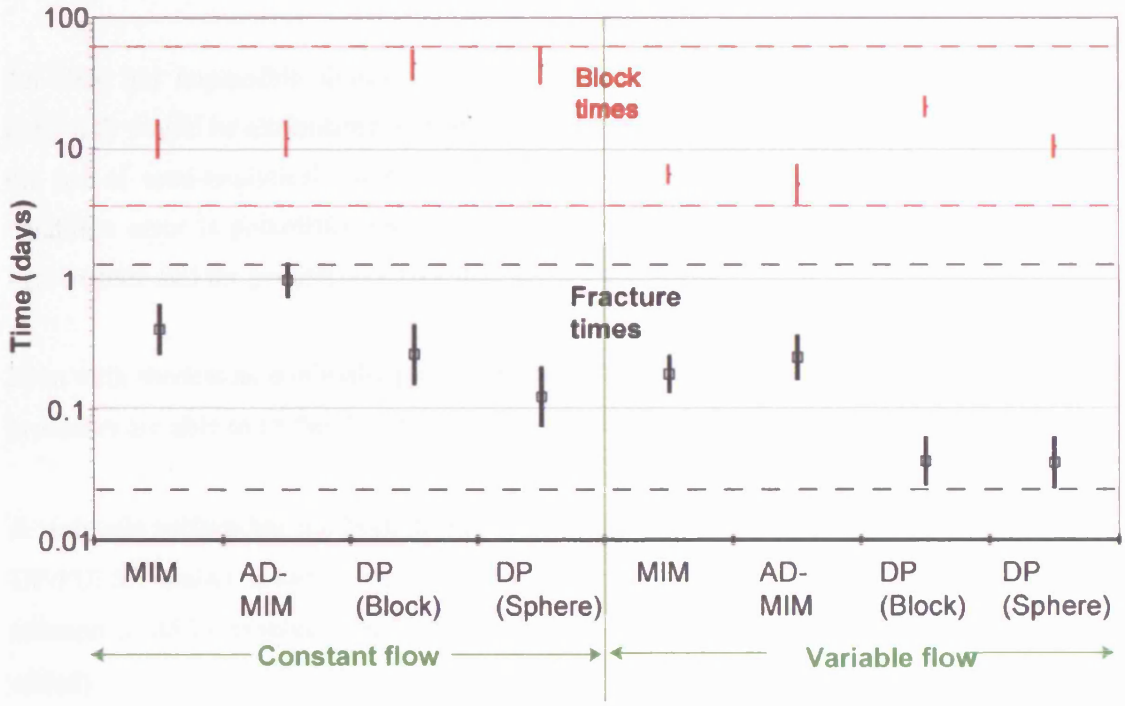


Figure 10.19: Range of characteristic times (95% CI).

10.7.2.3 AD-MIM

The combined AD-DP process was simulated, but was very slow to run and was not practical to use in conjunction with PEST. To understand the effect of the combined processes the simpler combination of AD-MIM was explored.

The calibration of the AD-MIM model fails to converge reliably and there is a correspondingly large CI for α. The starting condition (first guess) of parameters affect the calibration. In order to get a more meaningful calibration, the fitting was performed by successively fixing α and then fitting the MIM times. Plotting the correlation against α, revealed that the (weighted) fit improved very slightly as α decreased from ten to zero. For α >10 m, the fit worsens substantially. In summary, a slightly better fit results if no AD is assumed at all. However it is a sufficiently small improvement not to be a clear conclusion. The most demonstrable (but disappointing) result is that it is virtually impossible to resolve satisfactorily to what extent the AD process participates.

10.7.2.4 Discussion

It is clear from Figure 10.19 that there is wide range of characteristic times. It is likely that this will result in a wide range of predictions. Further work is therefore required to constrain both process diagnosis and parameter identification by development of more distinguishing tests.

An ideal but impossible situation for model comparison would be when deviations from a perfect fit would be attributable to inadequacies of the different conceptual models. Because of the use of semi-analytical models the numerical error is negligible. However, the boundary condition error is potentially rather large, as there was significant uncertainty over the flow against time and the porosity changed during the experiment.

Even with models as minimally parameterised as they are here, for certain conditions different processes are able to fit the BTC equally and reasonably well.

A residuals surface has not been generated for this data, as the multiple-flow rate correction to DP-PULSE makes it extremely slow to compute. However, in principle, whether a unique solution could be expected can be tested by investigating synthetic data (with or without noise added).

Whether a search algorithm will find the 'true' minimum depends on how it is conditioned – there is a risk that it could find different points in the valley containing the global minimum, depending on where the search is started.

The modelling work can be summarised as follows:

- Even for a relatively well-controlled tracer test it has proved difficult to identify the relative contribution of AD and DP processes.
- That notwithstanding, the block time, t_{cb}, was estimated to be of the order 6 to 20 days.
- No single DP geometry is identified - the MIM model works just as well as the fully diffusive model, so should perhaps be preferred, because it is simpler.
- Future experiments should be designed with careful consideration of heterogeneity and to be more diagnostic (see Chapter 8).

10.7.3 Heterogeneity

10.7.3.1 Multi-streamtube AD approach

Before the extent of the non-uniqueness difficulties were appreciated, an attempt was made to characterise the heterogeneity by examining the calibration of each outlet port.

How the flow rate was allocated required attention, since the flow profiles for individual ports differed considerably from the total flow pattern for all 13 ports combined. The same averaging 'stress' periods as were used for the total were adopted, but the constant flows were calculated for each port. This had the advantage that the flow discretisation pattern remained consistent for all 13 ports, but had the disadvantage that the averaging periods were no longer being selected to break the signal into a series of relatively constant periods of flow.

The AD and DP models were fitted to each individual port BTC assuming each flow-rate pattern. No discernable pattern emerged and it was not clear how to make practical use of the values. Most of the calibrations were taking long periods of time to complete (of the order of several hours on a PC), and none of the adjustments to the PEST control parameters suggested in Doherty (1994) in the event of convergence problems provided much improvement. The calibrations appeared to be suffering from a costly combination of non-uniqueness combined with slow-running codes. It quickly became obvious that the non-uniqueness issue would disrupt any pattern that may have emerged from analysis of the separate ports. Thirteen may also be an insufficiently small sample size, for revealing the shape of the parameter distribution. Given that at the same time no clear theoretical way had been developed which might handle parameters derived from separate port BTCs, the exercise was abandoned.

10.7.3.2 Conceptual progress

Although no clear recipe was developed for coping with individually calibrated port BTC data, some of the thinking that arose around the issue is worth capturing.

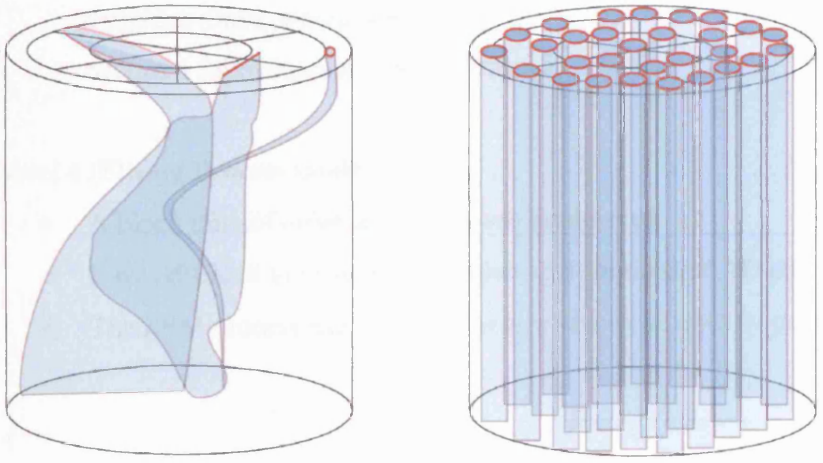
The issue of the scale of the sampling device (i.e. the outlet ports) affecting the results is an open question in hydrogeology[73]. There is a wide literature and debate on linking micro to macro scale observations. The discussion can reach to the heart of 'core' principles, such as the REV. Baveye and Sposito (1984) even suggested moving away from the concept of the REV and defining variables operationally.

How might knowledge of the 13 individually calibrated parameter sets be combined to predict what the parameter set for the total BTC might be? The answer depends on how each BTC is thought to relate to each other (is there lateral dispersion that causes the output of the ports to be correlated, or might each outlet be considered to be a separate streamtube?). In addition are

[73] Another example of this is encountered in this thesis is in Chapter 13, relating to how to interpret ingress to tunnels occurring at specific locations.

there extreme situations which can be compared and can similar distributions of parameters be predicted by statistical methods?

One possibility might be to assume a continuous distribution of stream-tubes, which are randomly distributed over the outlet area. Each outlet will sample a proportion of these tubes, scaled by the area that each outlet is apportioned. For example, Figure 10.20(a) shows a few large stream-tubes which are randomly intersected by the outlets, and Figure 10.20(b) shows a continuous distribution of non-correlated infinitesimal stream tubes and a question over whether they are correlated or laterally linked.



(a) (b)

Figure 10.20: A few dominant streamtubes (a) versus high density of stream-tubes (b).

10.8 *Conclusions*

This experiment enabled interesting observations, both in terms of what can be said about the system, but also what as yet cannot. Since it was the first experiment of its kind it has stimulated a lot of thought about how to improve the experiment and the modelling methodology. The questions posed by the data stimulated much of the theoretical discussion in Part I which is motivated by the desire to better diagnose processes and identify parameters more accurately. As a first step in an iterative research programme, this foundational dataset has proved to be invaluable.

In terms of what was found about the transport system it is worth splitting the analysis into a hierarchy:

Level 1 (Transport connection):

- Not all the outlet ports were flowing at once, but all flowed at some stage during the test.
- Both the average and the time-variable flow rates were highly variable between ports.

Level 2 (Key features):

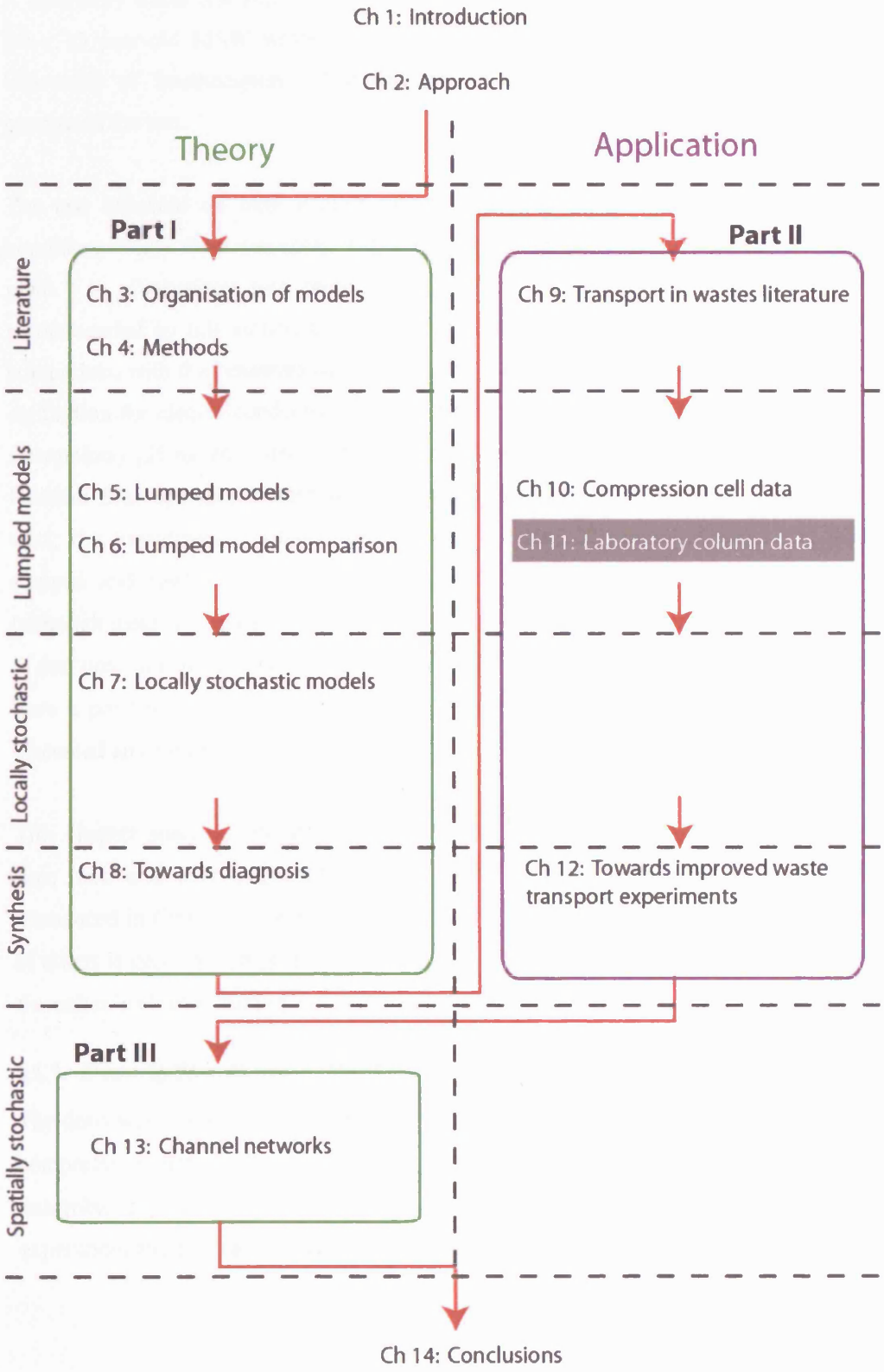
- Early time data were difficult to analyse conclusively.
- Concentrations measured at the outlets jumped during a period of no flow, probably indicative of DP exchange.
- The test produced a relatively well behaved monomodal BTC.

Level 3 (Statistics):

- The complicated input function meant that simple moments could not be established.
- The most useful statistic was the variability in concentration between the ports at different times, which was clearly correlated with the variability in total porosity that occurred following the flow disturbances.

Level 4 (Fitting Process models):

- A block time of order of 20 days was established
- It was difficult to resolve the relative contributions of AD and DP processes
- The MIM process was preferred as it produced an equally good fit to DP but is simpler.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

11 Laboratory column data

A laboratory tracer test was performed under saturated conditions (ponded inlet and raised outlets) for a 15 year-old MSW waste in the School of Civil Engineering and the Environment at the University of Southampton. Raw data was supplied by Layi Oni and Richard Beaven who conducted the test.

The test followed on from a prior experiment on the same waste with presumed unsaturated conditions where there was spray-irrigation, no ponding and outlets draining directly from the base-plate. A comparison was presented in Beaven et al (2005b). For a number of reasons, recommended by this author, the comparison remained primarily qualitative. Concerns about the comparison with the 'unsaturated' test included: lack of continuity between researchers resulting in calibration for electro-conductivity (EC) for the 'unsaturated' test (which was based on only two data-points) giving an order of magnitude difference to the saturated 'calibration' gradient; the Brilliant Blue dye (BB) absorbance wavelength was changed from 590 nm to 690 nm between the tests; the 'unsaturated' waste apparently had a higher permeability; the 'unsaturated' test was stopped and started, so that it may have required modelling by a transient-unsaturated code (although there is no data on whether any important saturation transients occurred, the water content of the 'unsaturated' test was not established, so all that was known about the system is that it did not have a ponding inlet condition). Because of these concerns this earlier unsaturated test is not discussed any further.

This chapter analyses data previously collected by colleagues. The data is analysed for the first time here and in a systematic fashion, by proceeding through the 'BTC analysis hierarchy' introduced in Chapter 4. Firstly, the basic conditions in the test are described (for which the work of others is described in Sections 11.3 and 11.2.2.2). Section 11.2.2.1 is review material. The work thereafter is all new analysis.

11.1 *Description of experiment*

The downward flow 'saturated' test was performed in an acrylic cylinder surrounded by an 'Eland' compression frame. The waste had not been disturbed from an earlier experiment investigating anaerobic digestion (Lucy Ivanova, personal communication). The key information about the experiment and the tracers are given in Boxes 11.1 and 11.2 respectively.

Figure 11.1: Photograph of laboratory-cell (from Richard Beaven).

Flow through the cell was achieved by application of a constant head across the cell due to a 'pond' of liquid above the waste and adjusting the level open to atmosphere of a set of inverted 'Y' connecters on the outlet pipes (the second half of the 'Ys' were connected to pipes that went to sample vessels). The pumped flow to the pond came from storage tanks containing either water-tracer mixture for the injection or just water for flushing. This recirculation was allowed to continue for seven days in order to achieve stability.

An approximation of a 'top hat' tracer injection pattern was achieved by pumping out the pond and replacing the pond with brilliant blue and sodium chloride tracers mixed in water pumped from a second tank. The continuous flow was then resumed, but with flow taken from the tank with tracer in it. At the end of the injection period, the pond was pumped out and replaced with water. However, the water was observed to immediately turn blue, so the pond water was changed a further three times (the whole procedure taking 45 minutes). Continuous flushing was then resumed, this time from a tank containing tap water.

No gas production was observed and the risk of entrapped air was minimised by first filling the waste from the bottom upwards (as opposed to filling from the top).

The gravel beneath the waste in turn sits on a geotextile membrane. Beneath this there are eight discharge ports on 45-degree angles at a radius of 120 mm. Allowing all eight ports to drain resulted in a very low flow in some of the ports, such that it was impractical to measure. Therefore it was decided to only use four ports (every second port) and to close the remaining four (Layi Oni, personal communication). Level 1 (transport connection) was therefore demonstrated for the ports that were used.

Box 11.1: Experimental set-up key data.

Waste type	Aged waste from Rainham landfill, Essex, subsequently shredded (max 40-50 mm), reconstituted and used as a matrix for anaerobic reaction of sewage sludge
Compression	47 kPa compression (released before tracer test with no rebound observed)
Dry density	0.572 t/m^3
Cell internal diameter	480 mm
Waste height	435 mm
Pond depth	154 mm
Bottom gravel	100 mm (30.7% water-filled porosity)
Hydraulic conductivity	1.06×10^{-6} - 1.76×10^{-6} m/s (range over 28 measurements)
Drainable porosity	7.5/8.8% (taken following downflow/ upflow)
Total porosity	Not recorded
Composition (dry weight)	Paper 9.6%, Plastic 16.7%, Textile 3.9%, Wood 3.7%, Metal 1.8%, Other 61%.

Box 11.2: Tracer test key data.

Tracers	Coomassie Brilliant Blue (measured as absorbance), NaCl (measured as EC)
Background BB* (leachate)	0.0089 Absorbance Units (AU)
Background BB (water)	0.003 Absorbance Units (AU)
Background EC (leachate)	1.48 µS/cm
Background EC (water)	0.63 µS/cm
Injected BB	3 g/L
Injected NaCl (EC)	10 g/L (EC=13.9 µS/cm)
Injection	27.25 L of tracers plus water injected over 14 days (with pauses)
Washout	291 L pumped over 24 days (with pauses)
Target flow rate	0.5 L/h
Total duration of test	52.25 days

*BB stands for Brilliant Blue dye

11.2 *Data analysis*

11.2.1 Flow rates (Level 1)

The flow was only recorded in four ports for 14.4 days, thereafter only the flow in port A was recorded (by using an automatic sampler which was not available at the start of the test). This meant that the flow in the other three ports needed to be estimated after this time in order to obtain a mass balance estimate.

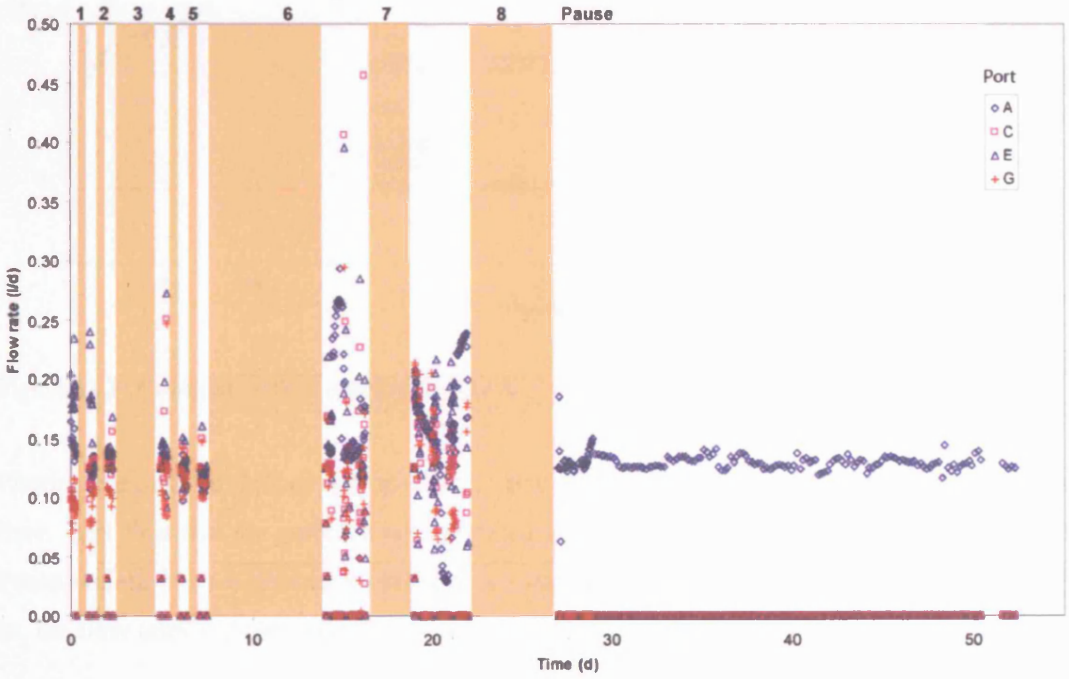


Figure 11.2: Instantaneous flow from each port, with pauses periods marked in.

Plotting the flow in each outlet (labelled A, C, E and G) revealed a noisy picture (Figure 11.2), although the noise reduced during the last 20 days or so, when there was no flow interruption. It is difficult to pick out any patterns from the noisy signals at earlier time. The pauses were at times when the apparatus could not be attended, due to a concern that any problems with the experiment might cause flooding in the laboratory.

Plotting the relative contributions of each port in Figure 11.3, shows that at around the 15 day point, the relative contributions from each had almost stabilised. The adjustment may have continued a little, but the relative rates at 15 days provide a reasonable basis for an extrapolation to estimate total flow.

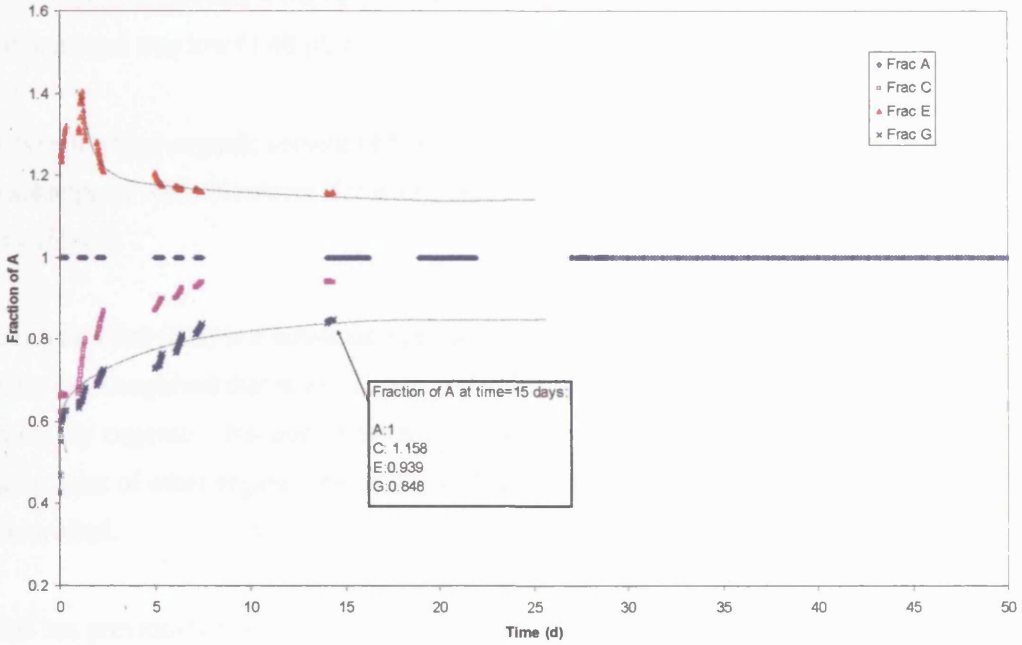


Figure 11.3: Cumulative flow relative to port A .

Plotting cumulative flow as a proportion of port A flow shows how the relative flows change with time. It is clear that the ports are not switching in the same dynamic pattern, as was the case for the Pitsea outlets, but are gradually settling to a constant relative rate. However, for the first 15 days or so, the flow rates and their relative contributions are still readjusting.

The flow is more tightly controlled at later time (port A flows at an average of 0.13 L/h $\pm$0.2 L/h from 15 days to the end of the test). The average for all ports based on the relative rates at 15 days is 0.52 L/h.

A simple ponding flow system provides adequate control for flow through aged waste, as long as the test is designed without any flow interruption and is allowed to settle down for long enough.

11.2.2 Tracers

11.2.2.1 Properties

Sodium Chloride (NaCl) has been widely used as a groundwater tracer and has the advantages that it is simple, cheap and easy to obtain, non-toxic and easy to detect using an EC meter. It can suffer the disadvantage that if used in too high concentrations it can affect density. It suffers where there

is a high background TDS (e.g. in MSW leachates). However, for this test in aged waste the background was low (1.48 μS/cm).

Given the high organic content of MSW, the cation exchange capacity may be sufficient for marked exchange of Na^+. However if it is EC, rather than Na^+ that is being measured, the EC BTC may be unaffected.

Brilliant Blue (BB) is a non-toxic dye tracer, traditionally used for marking flow-patterns in soils. It must be recognised that it will almost certainly preferentially become sorbed to and stain a certain – probably organic – fraction of the soil. Consequently the dye might usefully be used to mimic the properties of other organic chemicals, such as pesticides, but the analogy would have to be carefully controlled.

BB has previously been used as a tracer in wastes by Rosqvist and Bendz (1999). In theory there is no objection to using a sorbing dye for transport modelling in the context of a well-controlled laboratory experiment if the sorption properties of the dye in the same medium can be tightly constrained.

Flury and Flühler (1995) examined the tracer for use in soils. They fitted Freundlich isotherms for BB in a variety of soils. They also demonstrated in a field experiment that it is retarded in comparison to the arrival of Br^-, which they described as conservative in that particular soil. Rosqvist and Bendz (1999) used it in aged MSW waste, citing Flury and Flühler (1995) as supporting the assumption that the adsorption of BB can be assumed to be minor for tracer tests with high velocities, but it is unclear how this is justified. BB was used as a visual check on whether there was leakage along the lysimeter walls and was applied only to the periphery of the top surface of the waste. The arrival of BB led the Li^+ arrival, so it was concluded that there was preferential flow along the walls.

Stimulated by the need for better characterisation of BB in a waste substrate an MSc study was instigated (Jing-Lin, 2005). This was presented by Stegemann et al. (2006). Unfortunately without destroying the laboratory cell the waste in the cell could not be used, so a similar waste was sampled from Pitsea landfill[74]. Both a kinetic effect (order of six hours) as well as a non-linear isotherm were observed. The isotherm parameters depended on the sample size. Fitting a

[74] This was sampled, coned and quartered by this author with Richard Beaven.

Freundlich isotherm gave exponents $N = 0.38$ based on small (300 g) samples and $N = 0.77$ for an isotherm based on a larger sample (3000 g).

The BTC for a tracer with a Freundlich isotherm and $N < 1$ would be expected to have a sharper leading edge and a more dispersed trailing edge compared to a conservative or linearly retarded tracer (Chapter 5).

Depending on pH, BB may dissociate to a mono- or bi-valent anion. The anionic species can possibly form ion pairs with cations (e.g. Ca^{2+}), changing the sorption behaviour (Flury and Flühler, 1995). There was no literature found documenting decay of BB.

11.2.2.2 Instrument calibrations

The absorbance against concentration for BB was calibrated by adding BB to distilled water. BB was linearly related to absorbance for relatively low concentrations, but saturated at around 0.6 g/L and remained constantly saturated for concentrations above this. The experiment relied on the large attenuation since the input concentration of 3 g/L was above this saturation level.

The EC probe was calibrated in tap water, yielding a linear relationship between concentration and EC. The 'equilibrium' EC of the leachate when in recirculation was fairly low (1.48 μS/cm). The EC of the flushing water was lower still (0.63 μS/cm), so there was a change in 'background' concentration when the wash started.

Both calibrations were only performed once, at the start of the experiment, and it would have been desirable for them to have been repeated at the end to check for any instrument drift. However, the fact that the BTCs tended to low levels at the end suggests that if any error did creep in, it was not substantial.

11.2.3 Level 2 (Key features of the concentration data)

There are a few key comments relating to features of the BTC that require noting:

- Because of the flow interruption it is not straightforward to pick out many 'classical' features in the BTC. The 'jumps' in concentration following each interruption are a useful feature in their own right and probably indicative of DP exchange.

- Plotting on a cumulative flow axis (Figure 11.8) showed that the 'global' EC peak slightly lags the BB peak.
- The first arrival was 1.06 days from start of test (0.39 days of flowing).
- Because the concentration data spans less than three orders of magnitude, the late time gradient method cannot be used reliably.

11.2.4 Level 3 (Statistics: mass balance)

Mass balance was calculated as $\sum_{i=1}^{4}\int_{0}^{T} C_i(t)Q_i(t)dt$. In the absence of data after 15 days for three of the ports it was assumed that the flow rates at 15 days were maintained. Alternatively, allowing for variable flow, the product C_iQ_i is evaluated every 0.5 hours and then summed (again using the assumption of constant flow for the missing data). This is shown in Table 11.1.

Table 11.1: Mass balance results.

Tracer	Constant flow		Variable flow
	0.5 L/h	0.55 L/h	-
BB	24%	30%	21%
NaCl	101%	126%	84%

Assuming that only 75% of the BB injected was irreversibly sorbed (see Figure 11.5), the recovery for this 'mobile' proportion would have been for 96% 0.5 L/h, 120% for 0.55 L/h and 84% for variable flow.

The mass balance estimates vary considerably depending on the assumption of flow and on the assumption of how much of the BB is not permanently sequestered. The pond-changing procedure will have removed some tracer, but this would only have been from the very upper layers of waste which were in contact with the free water.

11.2.5 Level 3 (Statistics: moments)

Moments could not be taken because the flow rate was interrupted intermittently. Moments of the cumulative flow graph would have little obvious application or usefulness.

11.2.6 Level 3 (Statistics: concentration noise)

The EC and BB absorbance measurements were taken from all four ports for the full duration of the test. Figure 11.4 shows a comparison of the EC data from all four ports and the SD at the times when all four ports were simultaneously sampled.

The EC results are extremely close, with a relatively low SD between the four. For example for a point near the peak in concentration at 14.1 days the EC for port A is 9.15 μS/cm and the SD is 0.2615 μS/cm (i.e. SD approximately 3% of C). There is a crude visual correlation between the SD and the concentration of the BTC, which suggests noise increases with concentration. However, the peak of the noise does not correspond to the BTC peak and it appears that the noise is increased around the pause periods too.

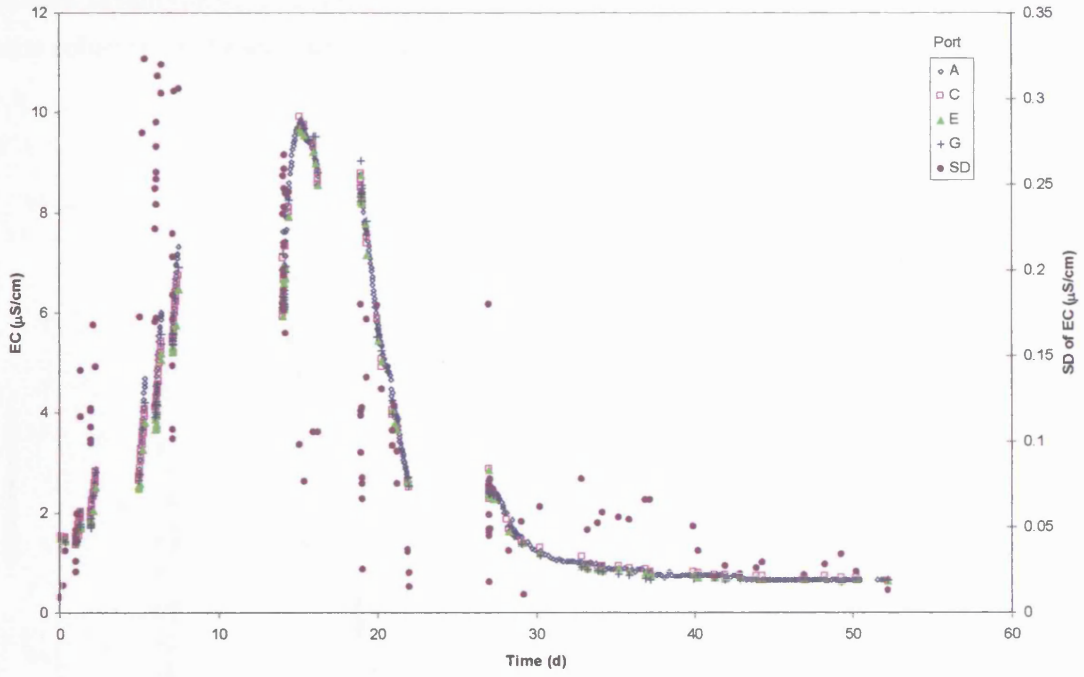


Figure 11.4: EC for each port (given in legend).

Figure 11.5 makes a comparison of concentration for the BB data. The SD of the data follows concentration, C, much more closely than for the EC and is a greater percentage of C. For example for a point near the peak in concentration at 14.4 days the BB for port A is 1.69 Absorbance Units (AU) and the SD is 0.29 AU (i.e. SD approximately 17% of C).

The precision of the instruments were EC 0.01 μS/cm and Absorbance 0.001 AU, meaning that the absorbance was more accurately determined at 14.4 days than the EC was. Tentatively the difference in SD appears to be due to physical causes rather than measurement (although it would be better to compare the accuracy by measuring the SD of measurements of a batch of a mixed standard than comparing instrument precision, but this has not been done).

Port A stands out as having a more 'peaky' response than the other ports. The difference between the noise levels in the two signals perhaps reflects an important difference. NaCl, if it can be assumed to be relatively conservative, demonstrates that the transport is homogeneous or laterally well spread. The increased noise in the BB signal may well reflect significant inhomogeneity of BB adsorption on to the waste. If this is the case and BB is required as a tracer, then characterisation of adsorption may need to take into account local adsorption variability at a scale commensurate with different stream-lines (i.e. at a smaller scale than a large 'REV' of waste that is a similar size to the entire column). A stream-lines based modelling approach may be appropriate in the future.

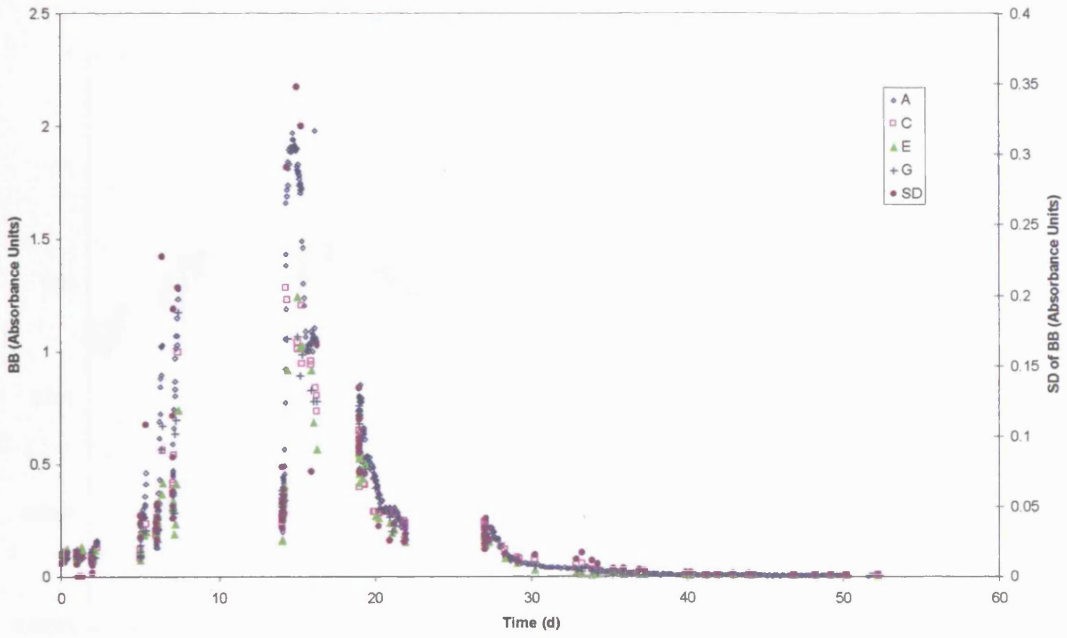


Figure 11.5: BB for each port (given in legend).

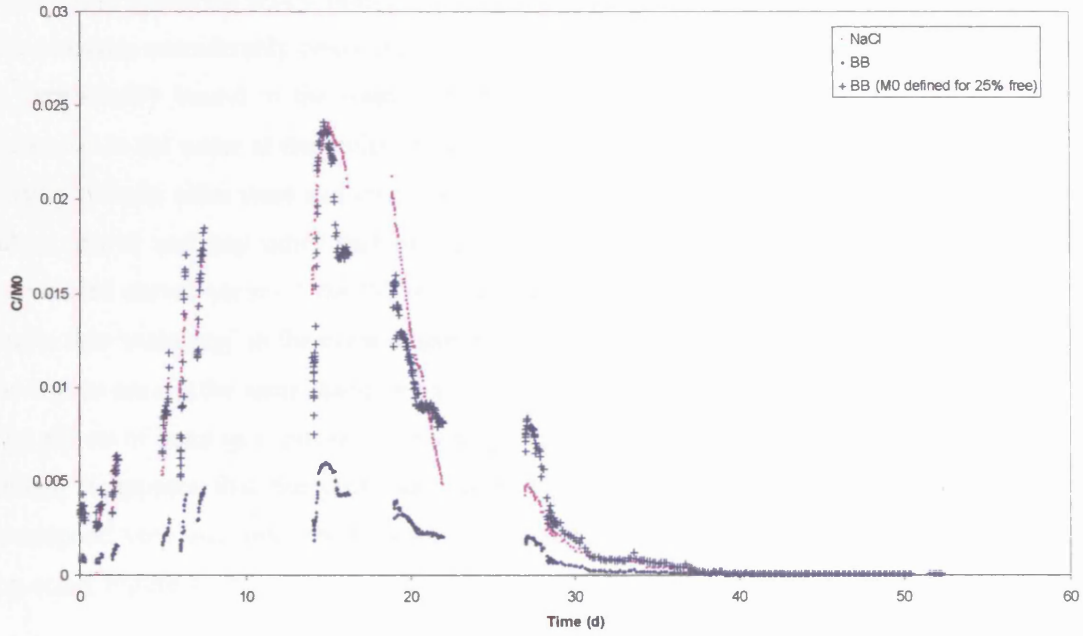


Figure 11.6: C/M_0 **for NaCl and BB for port A. Second BB plot is assuming input mass (**M_0**) is only 25% of what was injected in order to match the main peak.**

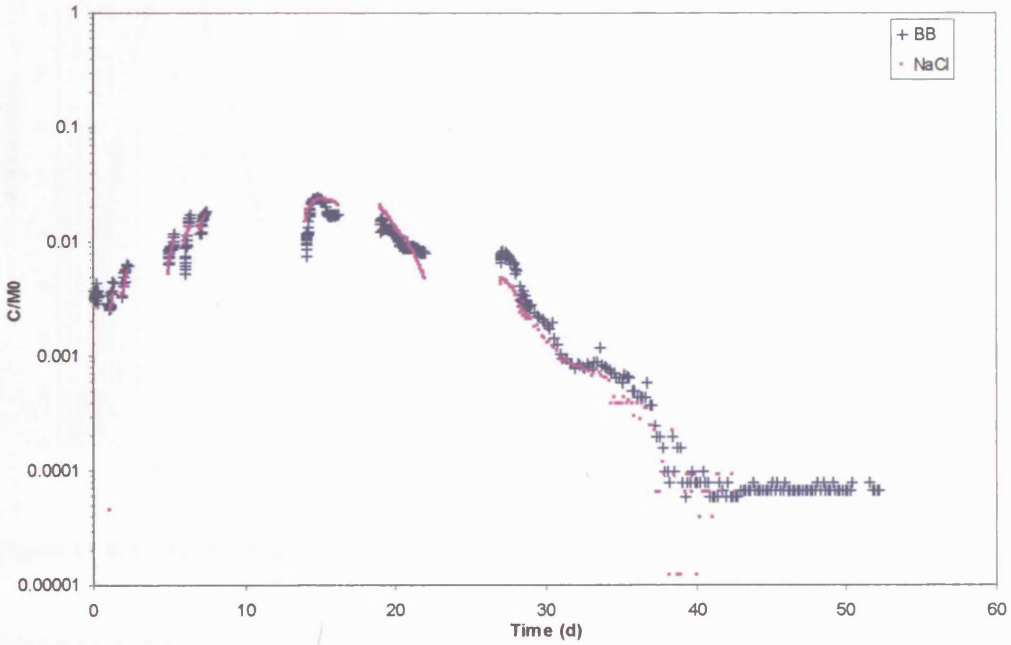


Figure 11.7: C/M_0 **for NaCl and BB for port A assuming** M_0 **is only 25% of what was injected in order to match the main peak (log-lin axes).**

Figure 11.6 shows the BTCs, normalised by the total mass of each tracer that was injected. The BB curve resides considerably below the NaCl one. A possible explanation is that a fraction of the BB is permanently bound to the waste. At the end of the experiment, when low absorbance was measured in the water at the outlet, the top surface and sides of the waste visible through the clear acrylic cylinder sides were still visibly stained blue. The amount absorbed by the cylinder, platens, tubes, gravel and any other part of the apparatus was not measured. Crude matching of the normalised curves occurs if the BB retained is assumed to be 75% of what was input. Figure 11.7 shows this 'matching' in the event of assuming that only 25% of the BB is mobile. It is notable that the curves are not the same shape, but it is difficult to pick out a clear pattern as to how they differ. The effects of heterogeneous non-linear sorption may have a fairly complex effect on the BB BTC, though it appears that the peaks are more strongly influenced than the late-time tails, which correspond very well until the background concentration is encountered (most clearly shown on a log-scale, Figure 11.7).

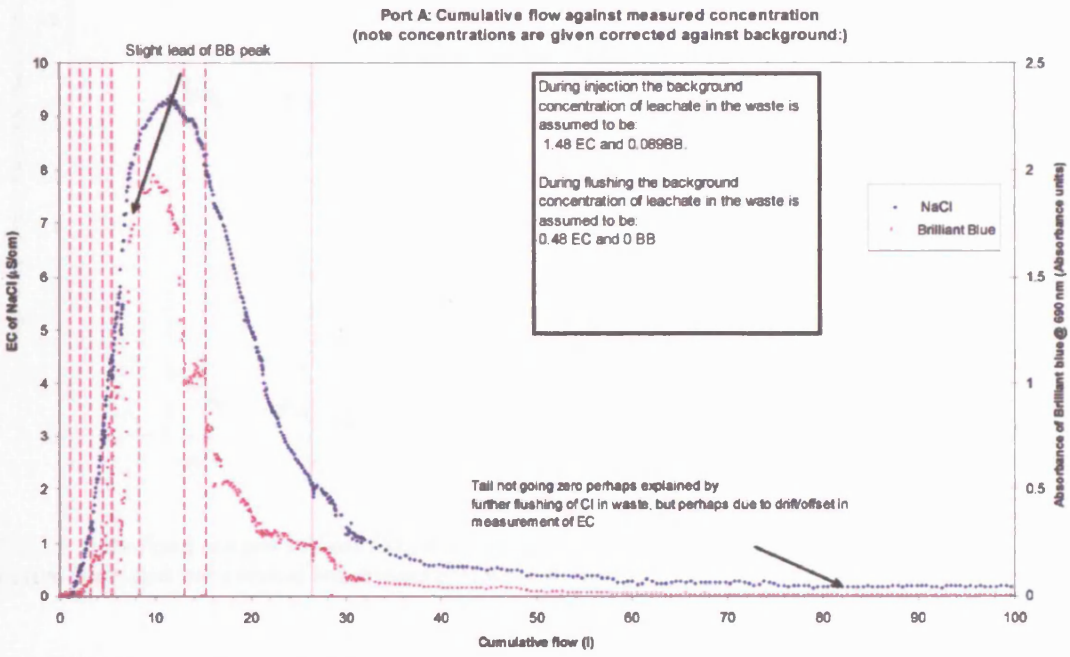


Figure 11.8: Concentrations for Port A, assuming 25% of BB is 'free' (vs. cumulative flow).

Figure 11.8 shows the BTCs in terms of the raw measurements for cumulative flow rather than time (but corrected for background). This compresses up the periods of paused flow, so that the 'jumps' in concentration after a pause appear as increased noise. The curves now look more like monomodal BTCs.

The peak for the BB occurs slightly before the EC peak, which at first appears non-intuitive as EC is not expected to be retarded and BB is expected to be extensively sorbed. Whether this observation is within noise requires verification (ideally a repeated demonstration). If it is verified it would raise the question of whether the EC curve is retarded, perhaps due to sorption of Na^+ or Cl^-. Stegemann et al. (2006) presented some limited evidence that Cl^- may be adsorbed by waste on a linear isotherm[75]. Aside from the peak, the rising limb of the BB does lag the EC.

'Tidying up' BTCs in this way provides a partial way of comparing two tests with different inlet flow patterns. Many waste studies use the liquid to solid ratio (LS) as the x-axis and some use cumulative flow, for example Rosqvist et al. (2004) used cumulative flow to compare the Pitsea data to the data from the in-situ lysimeter of Rosqvist and Bendz (1999).

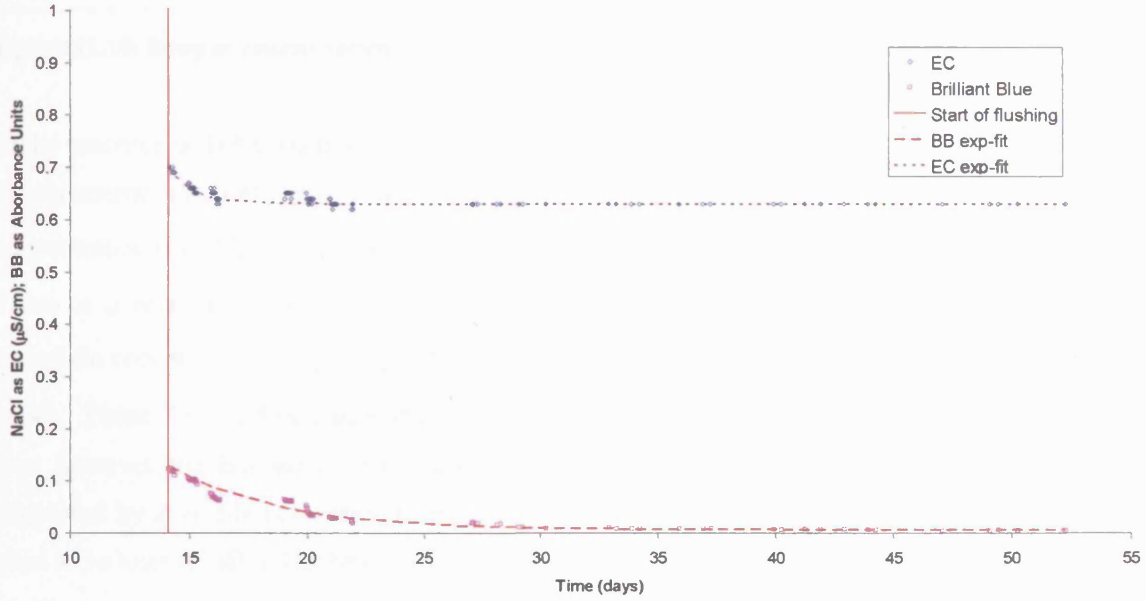


Figure 11.9: Pond concentrations following commencement of washout, with exponential best-fits as dotted lines and red vertical line denoting start of flushing.

The EC concentration in the pond (Figure 11.9) can be fitted very well with a simple exponential. The mixing constant for the pond is 8.8 days if it is assumed to be fully mixed, whereas the exponential is best-fit at 4.5 days, which is perhaps indicative of incomplete mixing in the pond.

[75] The 'isotherm' requires repeat measurement and experiments to demonstrate that the effect is not a function of background noise. Although Cl^- is generally assumed to be conservative it is still a charged particle subject to electrostatic interaction, so the possibility is not implausible. Stegemann et al. (2006) reported $K_d = 6.4$ L/kg which combined with the measured dry density 0.572×10^3 kg/m^3 and an assumed $\theta = 0.4$ would give a rather high $R = 10.2$, for a supposedly conservative ion.

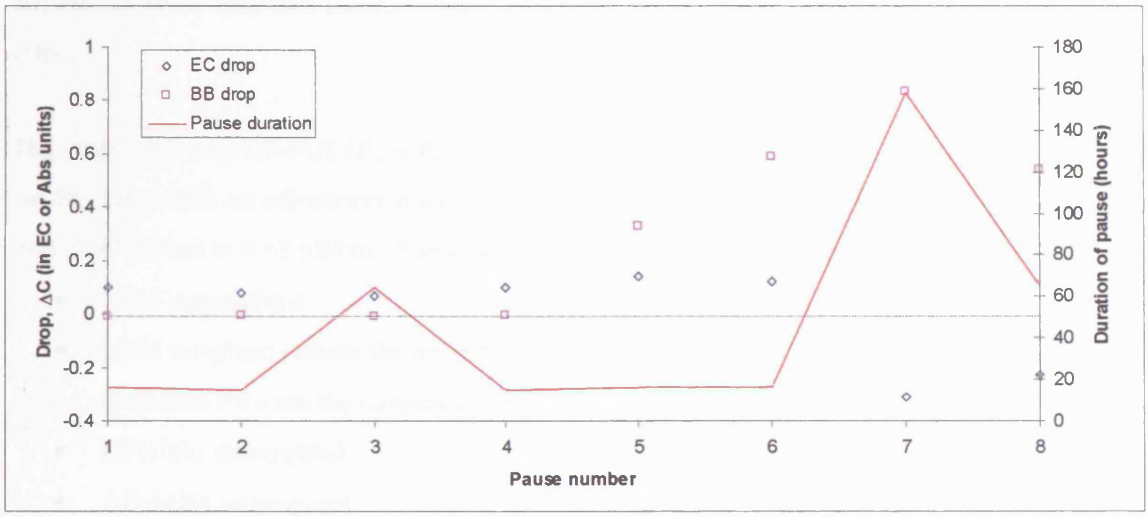


Figure 11.10: Drop in concentration (ΔC) at each pause.

In the presence of DP mass-transfer it would be expected that before the peak there would be a drop in concentration after each pause because the mobile concentration exceeds the immobile concentration (i.e. $C_m > C_{im}$), therefore there is a transfer of solute from mobile to immobile zones. There is a reversal some time after the peak, when the mobile concentration is less than the immobile concentration ($C_m < C_{im}$), therefore there is a transfer of solute from immobile to mobile zones. Pause 7 is the first pause after the peak. Figure 11.10 shows that EC profile does exactly this, however the BB increasingly drops. This is perhaps because BB is decaying. This is supported by a visible correlation between the later BB drops and the length of the pause. There need to be tests of BB in leachate to examine whether there is indeed any decay.

11.2.7 Level 4 (Modelling)

The modelling of the Pitsea data in the previous chapter showed that extensive fitting of the variable flow codes are time-consuming and not very revealing in light of the level of input uncertainty. The same general concerns apply to this laboratory data: flow rate was poorly controlled (at least for the first 15 days), incompletely measured and the tracers were poorly characterised. The large uncertainty surrounding the mass balance is a warning that fitting of models will be similarly uncertain. Hope of distinguishing between processes is almost certainly lost and any delineation of likely parameter ranges would need to be carefully handled to ensure the full uncertainty encapsulated in the concentration boundary condition (pond concentration), flow measurement and

that instrument calibration is carried through to the fitted parameters. In common with the Pitsea experiment (Chapter 10) the presence of some form of DP exchange is demonstrated, but it is difficult to know how this balances with other facets of the system such as mechanical dispersive effects.

The code used was DP-PULSE, with PEST as a fitting algorithm, using 50 comparison points for the EC data, with an adjustment made to compensate for the change in background concentration from 1.48 μS/cm to 0.63 μS/cm. Five cases were fitted ignoring the gravel:

- MIM unweighted.
- MIM weighted (where the weights at each time point are w=1/SD, where SD is the standard deviation between the concentrations measured in each of the four ports).
- DP(slab) unweighted.
- AD-MIM unweighted
- AD unweighted

Table 11.2: Best fit parameters for the three cases

Case	α / L	θ_m	t_{cb} or t_{IM} (days)	t_{cf} or t_M (days)	σ	r^2
MIM unweighted	-	0.06	0.51	9.1×10^{-3}	7.5	0.9941
MIM weighted (1/SD)	-	0.06	0.57	6.5×10^{-3}	9.4	0.9890 (weighted)
DP(slab) unweighted	-	0.04	1.59	1.3×10^{-2}	11.1	0.9944
AD-MIM unweighted	0.1	0.056	0.25	2.1×10^{-3}	11.0	0.9941
AD unweighted	0.17	0.25	-	-	-	0.9840

All the fits roughly mimic the peaks and slight fall-offs following pauses but all under-predicted the tail (Figure 11.11), something that is more clearly shown on a log-lin plot (Figure 11.12). The weighting did not noticeably improve the fit to the tail, despite increasing the importance of the tail in the fit. The calculated confidence intervals (CIs) were tightly bound around the best fit. The fits for the main peak appears visually convincing, although under-predicts the tail. The inclusion of the AD process slightly improves the tail. The AD process alone has the best-fit to the tail, although clearly cannot model the jumps in concentration at the pauses.

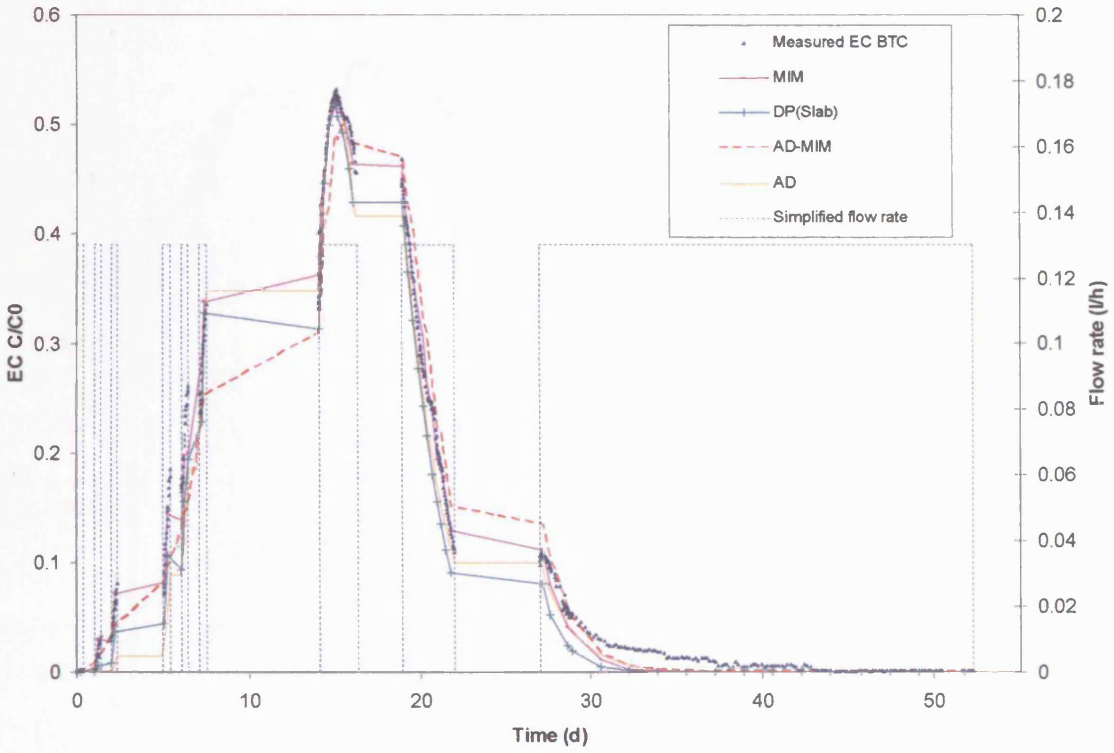


Figure 11.11: Single fit to the laboratory EC data using DP-PULSE (MIM), parameters given in Table 11.2.

The unweighted MIM fit found that mobile porosity, $\theta_m = 5.7\%$, which is not far off drainable porosity, $\theta_d = 8.9\%$. The porosity ratio $\sigma = 7.5$ gives immobile porosity, $\theta_{im} = 42.8\%$ and total porosity, $\theta = 48.5\%$. For fitting to a DP(slab) $\theta_m = 4.0\%$, giving $\theta_{im} = 44.4\%$ and $\theta = 48.4\%$.

The characteristic block time, t_{cb}, was around one day. This is an interesting in contrast to t_{cb} of around 20 days calibrated in the Pitsea cell tracer experiment (Chapter 10). This difference might be attributable to the fact that the laboratory waste was shredded whereas the Pitsea waste was not, and therefore the block sizes in the laboratory waste were smaller. Alternatively it might be the case that the column size and length of experiment affect the apparent t_{cb}. A comparison between experiments at different scales with identical wastes would distinguish these two possibilities. It is worth noting that such a comparison was made for two MSWi bottom ash chloride flushing datasets in Beaven et al. (2005a). One dataset was based on flushing from a small lysimeter, the other from flushing from a pilot landfill. This lends weight to the hypothesis that experimental scale may have an important influence on the perceived timescales.

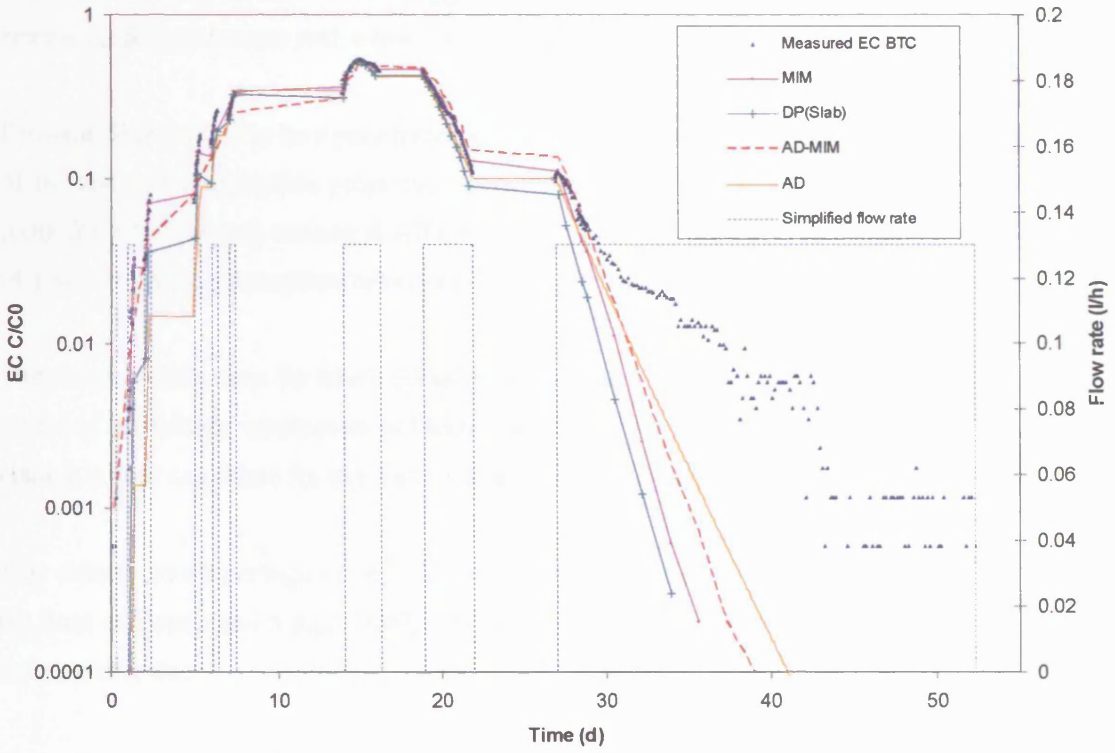


Figure 11.12: Fit to EC data using DP-PULSE (MIM) (log-lin axes, parameters given in Table 11.2).

Error will have arisen from flow rate approximation, inaccurate boundary conditions (inlet and outlet) and the fact that the MIM model is possibly inappropriate. Some error is introduced by the correction to the change in background concentrations and by the fall in pond concentration during the flush not exhibiting a perfect 'top-hat'. There may also have been some exchange or mixing in the basal gravel.

The design of the experiment, with many periods of zero flow, a relatively short column of material, the existence of a large top pond and a substantial volume of basal gravel all encouraged poor control and knowledge of the effect of the boundary conditions. This laboratory-test can (should) be relatively straightforward to improve and duplicate.

11.3 *Conclusions*

In terms of the control of the flow rate in the cell, away from periods of intermittent flow, the flow rate was reasonably controlled and showed little sign of long-term drift. Flow control by pumping to a pond appears to be feasible in an aged waste but requires a few weeks to stabilise.

The use of the cheap tracer NaCl, logged by EC as a proxy for concentration, appears to be very promising for aged waste with a low-strength leachate.

Brilliant Blue proved to be a poor tracer, either degrading or exhibiting sorption kinetics at the scale of the test. The adsorption properties of the tracer were not properly characterised for this waste (only for Pitsea waste), making it difficult to model the BB breakthrough without making a number of assumptions about sorption behaviour.

The characteristic time for block diffusion was found to be of the order of one day. Beyond this 'order of magnitude' assessment of block time, the poor mass balance and unquantified input errors make it a poor candidate for any further detailed modelling investigation.

This dataset could perhaps be revisited with a model that could handle advective transfers (along the lines of Bendz and Singh (1999)). However, it would be better to also improve the experiment before doing this.

No further modelling is justified for this data, as is expected to yield diminishing returns (the AD model could be used, for example). The experiment and data are best thought of as a pilot for developing further highly controlled laboratory experiments on wastes. The parameters can be used as prior values for improving experimental design. In terms of what was learnt about the transport system, the value of this data is mainly at the bottom of the tracer hierarchy (Table 4.1), as follows:

Level 1 (Transport connection)

- Only four out of eight outlets flowed. Because the four non-flowing outlets were subsequently valved-off, it is not possible to say whether they would have restarted intermittently or not.
- The flows in the identical and equally spaced four ports took around 15 days to settle. The flow rates varied from 1.16 to 0.85 times the flow in port A.

Level 2 (Key features)

- Early time was not very revealing.
- EC peak possibly lags BB (this may be due to decay or permanent removal of BB).
- There is an insufficient period of flushing and too much noise to reliably establish any late-time gradient.

Level 3 (Statistics)

- Moments are complicated by intermittent flow.
- The mass recovered was estimated as 84% for NaCl as EC and 21% for BB, when a variable flow rate was assumed.
- The variability in concentration between the ports can be used as a basis for weighting of fitting. The SD for the BB BTC is greater than for the EC BTC, perhaps explained by heterogeneous sorption. For both BTCs a relative noise appears most plausible, although it may be more complicated than that.

Level 4 (Fitting process models)

- Least squares fitting does not represent the tails well.
- It is not possible to distinguish between processes on the basis of goodness-of-fit.

This concludes the analysis of tracer data in wastes. For complete symmetry of analysis of theory and application there would now follow a section on locally stochastic models applied to wastes.

A multi-streamtube model was used in Beaven et al. (2005a) in order to take into account the geometry of landfills (e.g. the shape of the base of the landfill and mounding on the top surface). Transport was assumed to occur in 1D vertical stream-tubes. The average concentration at the base was estimated by integrating stream-tubes over the distribution of depths (e.g. using $\langle C \rangle = \int_0^\infty f(t,z)p_z(z)dz$, where $f(t,z)$ is the transfer function of a single stream-tube and $p_z(z)$ is the depth distribution, for example in Figure 11.13). This may prove to be a highly effective way of implementing simple 1D models to more complex field environments without the unnecessary complexity of building a full 3D model.

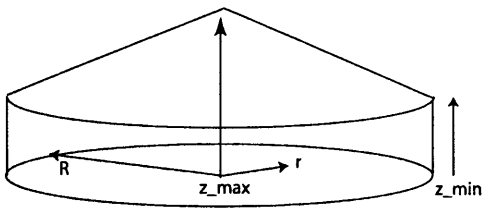
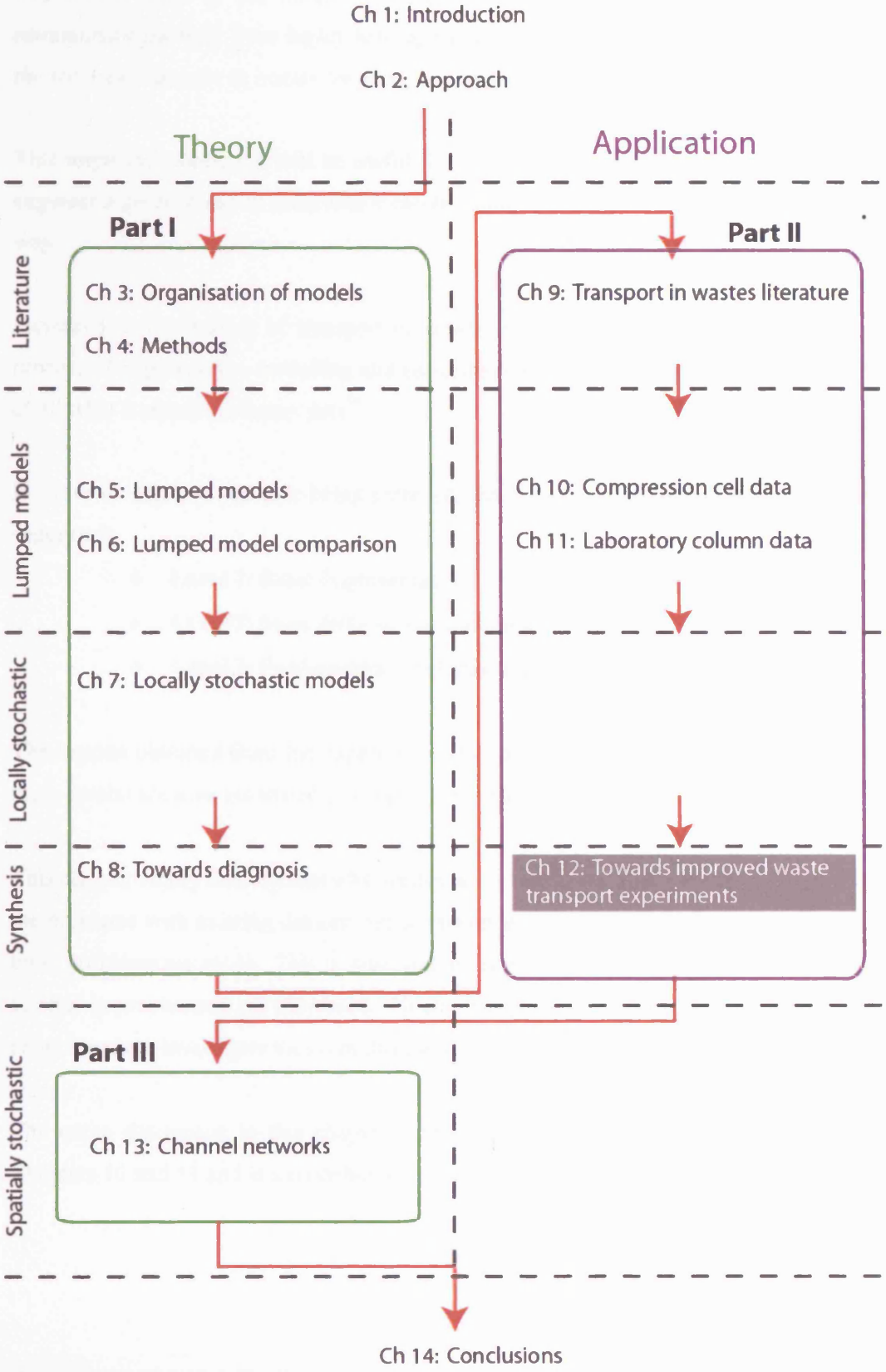


Figure 11.13: Idealised depth distribution of Vestskoven landfill. From Beaven et al. (2005a: Figure 2).

An attempt at a multi-streamtube model by analysis of the BTC at each port in the compression cell data in Chapter 10 did not yield any clear distributions of parameters. So, although these models

are likely to be useful in the future, the datasets explored here do not support their use (or indeed the use of spatially stochastic models) in calibration.

It therefore only remains in Part II to cover the 'synthesis' issue: that of how to improve waste experiments based on the earlier experience with data.


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Lumped models
Locally stochastic
Synthesis
Spatially stochastic
Ch 3: Organisation of models
Ch 4: Methods
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 7: Locally stochastic models
Ch 8: Towards diagnosis
Ch 9: Transport in wastes literature
Ch 10: Compression cell data
Ch 11: Laboratory column data
Ch 12: Towards improved waste transport experiments
Part III
Ch 13: Channel networks
Ch 14: Conclusions

12 Towards improved waste transport experiments

The central aims of the thesis are long term desires, they are: *to improve prediction of contaminant flushing from highly heterogeneous materials and to apply this generic desire to the study of transport in wastes (in particular in MSW and MSWi wastes).*

This improved prediction will be useful for making management decisions about how to best engineer a given waste in order that it can be brought to a low-pollution risk state in an optimal way.

Increased understanding of transport processes in wastes is likely to arise from an iterative process of experiments, modelling and suggested enhancements and further experiments. First of all what is needed is *better* data[76].

In order to improve the data being gathered, three levels of modifications to experiments are suggested:

- Level 1: Basic improvements.
- Level 2: More difficult and less obvious refinement.
- Level 3: Fundamental development of the experiment.

The lessons obtained from the experiments and modelling in the previous chapters (at each of these levels) are now combined to suggest completely new future experiments.

This chapter firstly summarises why wastes are inherently difficult media to work with. Next, the problems with existing datasets are discussed and specific suggestions as to how to improve these problems are made. This is structured in terms of the above levels of improvement. Next, general improvements are discussed. Finally, the primary questions for the future and possible experiments to investigate them are discussed.

The entire discussion in this chapter is founded on the challenges posed by the analysis in Chapters 10 and 11 and is a contribution to improving waste transport experiments.

[76] It is not always the amount of data, but what data and how it is taken that is important. Increases in data collection may not automatically yield commensurately better understanding of the system, or indeed better-constrained models. Berkowitz (2002:878) warned that, "too many (expensive and time-consuming) field measurements are made without due regard to the complexity of the system under consideration".

12.1 *Difficulties*

12.1.1 Difficulties specific to wastes

Firstly, the difficulty with analysing wastes is that there is a lack of useful taxonomy, in terms of the transport behaviour:

- A waste type (e.g. MSW) is not necessarily a single system since there are many different potential mixtures of what might be emplaced. It is therefore difficult to develop generic conceptual models of flow and transport.
- There is no clear classification beyond the broad type of waste and the basic pre-treatment (if any) and there is no classification based on the structure and condition of wastes *in situ*.
- There has been relatively little field characterisation of wastes.

Secondly, wastes suffer a series of general problems:

- There is a wide variety of chemical interactions occurring in both solid and liquid phases.
- Boundary conditions (water or leachate input, temperature, placement of waste, overburden, water levels) may change but are often inadequately measured.
- The filling of a landfill may take considerable time, such that the emplacement process itself may need to be a data input and incorporated in some models.
- Wastes may contain other materials as part of the site management, for example daily cover, a final cap, temporary roadways and gas and leachate abstraction infrastructure.
- It is desirable to make quantitative statements about waste heterogeneity and be able to relate this to other classes of geological media (especially soils and aquifers).
- Containment materials are difficult to categorise as they may age at timescales longer than is practical to measure. The top cover and daily cover introduce soil materials into the waste. Clays in particular are difficult materials to characterise due again to a dynamic structure (which may change as a function of TDS, pH and even the particular molecules sorbed). Clays have low hydraulic diffusivities so may be difficult to stabilise even at the laboratory scale. In common with wastes, a laboratory-scale sample of clay may completely miss heterogeneity scales. Cracks may be very important to transport at the field scale but unmeasured in the laboratory (Jørgensen et al., 1998).

Thirdly, there are difficulties specifically associated with organic wastes:

- Bio-degradation of wastes with organic content can cause generation of gas, heat and degradation products as well as structural change; i.e. a series of complex coupled processes.

- Bio-degradation may control pH and Eh, therefore affecting solubility and sorption behaviour of some solutes.
- Humics in leachate may play a role in colloidal transport.
- High strength (TDS, TOC) leachates cause difficulties with detecting tracers above background and in analysis.

Forthly, inorganic wastes present their own problems:

- There may be very (exceptionally) high concentrations of solute initially. MSWi bottom and fly ashes have particularly high concentrations of solute. High concentrations may mean that different chemical processes would need to be modelled and this can present analytical challenges (for example, at the simplest level, the relationship between EC and TDS becomes non-linear for high concentrations). Basic fluid properties such as density and viscosity are affected, too.
- Chemical changes may alter the structural and flow properties of the system with time.

12.1.2 Response to difficulties

The difficulty with the medium, the containment, limited data and conceptually restricted models combine to limit the reliability of prediction of solute flushing. The appropriate response to this combination of issues should neither lump all these difficulties together, leading to an acceptance that poor data quality is inevitable, nor accept that unenlightening empirical models are the most reasonable parsimonious representations commensurate with the data. Instead, arguably a much more systematic approach is needed to categorise and identify conditions under which certain processes or key combinations of processes dominate. Each problem should be addressed individually to build a more coherent understanding of the uncertainties and errors entering the system, of which the most important issues can be prioritised.

12.2 *Problems and suggested improvements for existing datasets*

12.2.1 Literature

Very few of the datasets in the literature appear to be useful for transport modelling purposes. The main difficulties with using old data are that the initial and boundary conditions of the system are not always established or reported, and that the data may have been collected with the intention of parameterising a simple model (such as the exponential model[77]). Results from

[77] If the exponential 'mixing' model is used, it may be that the cumulative leachate volume is reported but the flow rate and time (which are very important for DP processes) are not reported. An 'average' constant flow rate might be assumed, but to what extent this resembles the conditions in the system would remain unknown.

the European Standard for a waste upflow percolation test, CEN/TS 14405, (see http://www.standardsdirect.org/) are useful for comparison between wastes but the downside of standardisation is that all the tests suffer from the same weaknesses in experimental design .

12.2.2 Southampton laboratory experiment

The laboratory experiment described in Chapter 11 is now discussed in terms of improvements at each 'Level'. The trivial issues in Level 1 are perhaps too often glossed over and some may be more prevalent in Waste Science than is openly admitted. However, they should not be forgotten and are dealt with here. Failure at this level can seriously compromise the value of the entire experiment.

Level 1: Basic improvements

- Simple written procedures are needed to standardise methods to ensure comparability between different datasets.
- A standard datasheet would ensure no key measurements are overlooked.
- Having the same person collecting the data, rather than having a change would remove any bias arising due to the experimentalist.
- Avoiding pre-processing of raw data in the laboratory would eliminate the wasted time correcting and checking calculations done in the laboratory.
- Continuous data logging would ensure more regular and reliable sampling than manual sampling.
- Ensuring that no changes are made to an experiment that compromise a mass balance estimate is essential.
- For instantaneous measurement devices (such as absorbance, EC and ion-specific electrodes) the calibrations of the instrument at the start and finish of experiment should be compared and the measurements independently verified.
- The cell should be continuously weighed. This will allow the time variability of porosity to be determined. This is particularly useful for unsaturated investigation.

Level 2: More difficult refinement

Tracers

- BB should be abandoned as a cheap 'BTC' tracer. However, it is useful for marking flow paths and visualising edge-flow and there might be a case to use it to mimic certain organic contaminants (but considerable care would be needed over making this analogy).
- Other tracers should be used, such as deuterium and lithium bromide.

Input function

- Input should be a 'top hat' function of tracer into the upstream 'pond', which is automatically and continuously stirred. The input function should then be known with a high degree of accuracy. Although the improvement in the parameter CIs due to the better knowledge of the input function is difficult to quantify, what can be quantified is the effect on the CI of adding a mixer at the inlet to a column. An analysis of this effect for the AD model is given in Box 12.1.

Box 12.1: Effect of mixer on confidence intervals of AD model.

Compare the CIs for V and α for the AD model with different mixer sizes with a step-input of tracer to an initially empty mixer and column. The mixer has a time constant of $\lambda = Q/Vol$, where Q is the flow rate and Vol is the volume of mixer. The CI is found from the Hessian method for an increasingly long time (i.e. for more data points). The time at which the comparison is stopped is called the 'truncation time'.

The concentration due to input of $C(0,t) = C_0(1-e^{-\lambda t})$ is given by van Genuchten and Alves (1982: Equation A9):

$$C = C_0 e^{-\lambda t}\left[\frac{1}{2}e^{\frac{(V-y)z}{2D}}\operatorname{erfc}\left(\frac{z-yt}{2\sqrt{Dt}}\right)+\frac{1}{2}e^{\frac{(V+y)z}{2D}}\operatorname{erfc}\left(\frac{z+yt}{2\sqrt{Dt}}\right)\right] \qquad (12.1)$$

where $y = V\left(1-\frac{4\lambda}{V^2}\right)^{1/2}$

Mindful of the change of scales, compare the development of the CI shown in Figure 12.1 as the truncation time increases. Fixed parameters are Δt =0.1, $z = 50$, $V = 10$ and $\sigma_N = 0.01C_0$.

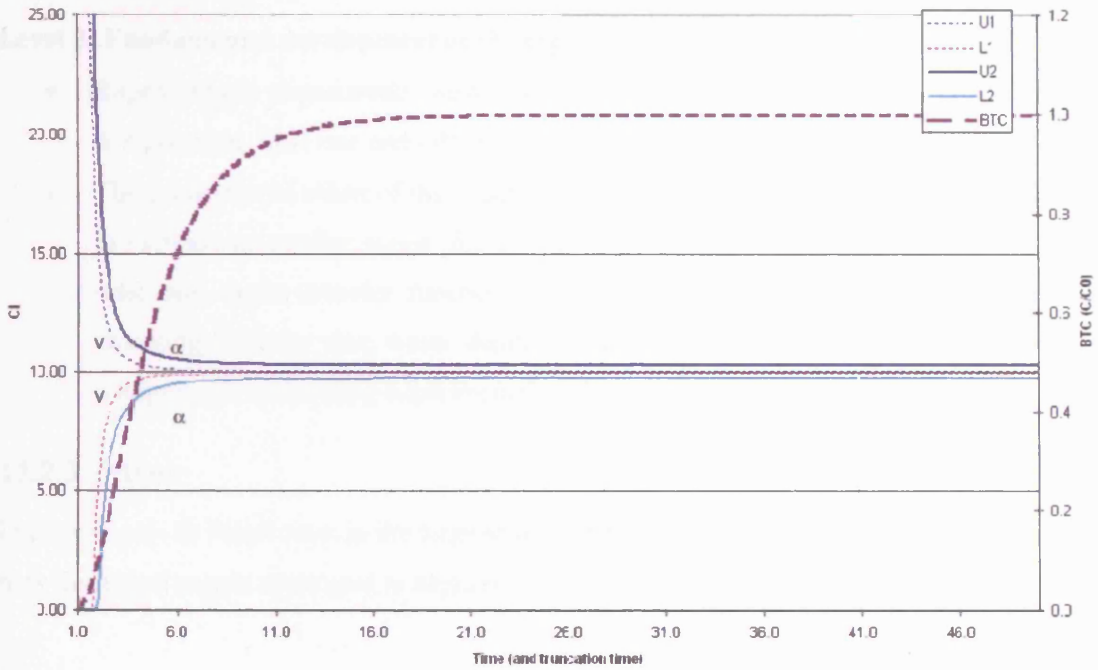


Figure 12.1: 95% CI for PARAM1=V **and** PARAM2=α **, unmixed (i.e.** $\lambda = \infty$**).**

From Table 12.1 it is clear that the bigger the mixing cell, the wider the CIs (most especially of α). Table 12.1 compares the CI (which is essentially stabilised at a constant value) at t =50.

Table 12.1: Summary of effect on the CI of changing the decay constant of the inlet mixer.

λ	**95% CI for** V	**95% CI for** α **at time=50**
∞	0.11	0.55
0.2	0.14	1.04
0.1	0.17	1.62
0.01	0.45	12.12

The non-mixed case ($\lambda = Q/0 = \infty$) gives the tightest CI. It is not entirely clear why this is the case. One explanation is that the delay caused by the mixer means that not all of the 'information' has had time to appear. This is not reasonable, since the $\lambda = 0.01$ case has been repeated for five times as long as the standard case and still tends to a relatively wide CI. A possible explanation may be found by appealing to the frequency content of the input; the spectrum for a top-hat contains a wider range of frequencies than for a more rounded, mixed, input function. A future question would be to quantify the effect on the CI due to the reduction in input uncertainty due to utilising a mixer.

Level 3: Fundamental development of the experiment

- Repeat tracer experiments with systematic changes in key process conditions (e.g. compression, flow rate and different levels of saturation).
- The proportional effect of the experimental apparatus on estimated parameters is likely to increase at smaller scales (for example, the effects of the side-wall, input function, inlet and outlet transfer functions). This could be systematically investigated by changing cylinder size, waste depth and gravel layers in order to make an informed compromise in deciding experimental scale.

12.2.3 Pitsea

Improvements to future tests in the large-scale compression cell test described in Chapter 11 are now discussed (again structured in terms of 'Levels' of improvement).

Level 1: Basic improvements

- The amount of gravel should be minimised and the waste volume maximised, in order to minimise the disturbance to the BTC from the inlet and outlets of the cell.
- An identical load-cell with known weight should be added, to act as a control for the four load cells beneath the compression cell.
- Equally-spaced equal-volume samples should be collected automatically. This has the added advantage of requiring less manpower.
- A more rigorous approach to quality control of analytical measurement is needed, through randomisation of samples and inclusion of repeats. Establish total measurement error over several measurements in addition to establishing instrument precision.

Level 2: More difficult refinement

- It is desirable but difficult to overcome the analytical difficulties of measuring bromide following acid digestion to release the full potential of the LiBr tracer.
- The use of a further tracer, perhaps deuterium, would enhance the transport information available from each 'run'.
- Simultaneously measure flushing of one or more components in the emplaced leachate.
- Gas emissions could be continuously recorded during the tracer experiment and would be a useful further measurement of the dynamics of the waste-fluid system.
- Consideration could be given to removing the peripheral port measurement from the dataset (there is strong evidence of preferential flow in the outer annulus nearest the wall).

- Physical measures to avoid or reduce the enhanced peripheral flow could be trialled.

Level 3: Fundamental development of the experiment

- It is desirable to achieve a constant total stress on the top of the Pitsea cell for all times during tracer tests. This is categorised as a Level 3 issue because it potentially involves substantially redesigning the sealing system for the upper platen, which at present is likely to leak if there is any movement once the seals are formed. In addition, further complications will be introduced as increased settlement will lead to increased compaction of the waste and hence a change to any hydraulic equilibrium in the cell. Operation of the cell at constant volume (as is undertaken now) avoids some of these issues. However achieving a control on total stress would also be desirable for geo-mechanics experiments.
- It would be worth aiming to measure transport properties only over a relatively short 'snap-shot', before the system changes (145 days is arguably too long for a fresh waste which will change substantially during this period). This might involve use of 'super-fast' tracer tests (as demonstrated theoretically in Chapter 5).

12.2.4 Improved approaches (all scales)

A series of possible future approaches to waste experimentation at all scales are suggested in this Section (effectively 'Level 3' improvements).

- There should be investigation of EC as a concentration proxy in well-degraded wastes (checking for Cation exchange).
- EC should be measured as a cheap and easy proxy for concentration. This gives an immediately accessible field reading of the state of the tracer test.
- It would be useful to conduct tracer tests over a range of flow-rates, right down to rates similar to current field percolation rates. Higher rates are also useful to evaluate possible improvements to flushing.
- Consideration should be given to the possibilities offered by destructive after-test investigations (e.g. adding a charge of dye-tracer and peeling back the surface layer by layer to investigate flow paths).
- Measuring pre- and post-test degradable content would establish how far the degradation has progressed during a tracer test.
- Systematic variability of the effect on transport of compression, water content and flow rate conditions for the same waste would give a very useful indication of the potential key engineering controls on flushing.
- Consideration could be given to suppression of bacterial decay (say by low temperature water or acidity). Otherwise, it may be worth applying more focus on aged wastes to

get the fundamentals of transport properly understood away from the additional complexity and potential lack of control caused by biological decay.

- The role of the top layer of waste in conditioning the noise level and flow rate of the various outlets needs to be investigated. This should test whether the difference in flow between the outlet ports is dominantly a localised outlet effect due to 'choking' of the outlet flow by the top layer or whether there are heterogeneities that cause whole stream-tubes to differ substantially. The simplest way to test this would be to dig off the top 10 cm and re-measure the flow in all the ports, perhaps then repeating with a further 10 cm and so on.

Handling the system changing with time

- 'Snapshot' tracer tests lasting much less time than the rate at which the system changes need to be considered in order to characterise transport behaviour of a waste at different points in time as it degrades.
- Transient saturation models need to be developed to handle variable leachate or water input (due to lack of control, intentional variability or for use in the field).

Approach to heterogeneity

- There is a need to move on from the considerable limitations of existing approaches, such as using an empirical mixing model (e.g. Landsim), black-box models (Maloszewski et al., 1995), or invoking a 'hetereogeneity parameter' (Zacharof and Butler, 2004), which is just a mechanical dispersion parameter.
- Spatially stochastic models are potentially very useful, even if they are highly speculative and based on limited geostatistics. Insights may come from observing how simple but reasonable underlying rules may give rise to 'emergent' behaviours which are similar to those recognised in real datasets. It is worth further investigating such models (e.g. channel networks such as those examined in Part III of this thesis).
- Once data is of sufficient quality to reliably work at 'Level 4' in the BTC analysis hierarchy it is probably necessary to maintain a wide range of models concomitant with the level of conceptual uncertainty. Fixating on one or two models may greatly weaken any attempt to explore predictive uncertainty.
- The effect of parameter and model uncertainty needs to be carefully handled. There have been promising and relatively recent developments in this direction. For example, pseudo Bayesian methods such as SUFI (Abbaspour et al., 1997), SUFI-2 (Abbaspour et al., 2004) and GLUE (Beven and Freer, 2001), other methods such as Pareto optimality (Gupta et al., 1998), as well as more formal Bayesian approaches (Thiemann

et al., 2001) are all in the recent literature. The USGS have recently released the 'JUPITER API', which is potentially a useful suite of tools[78].

- Experimental methods need to be modified in order to reduce the problems that stem from heterogeneity. A particular focus needs to be placed on the nature of measurement – the scale of averaging over which properties are measured (e.g. the area of the outlet over which fluid is measured, the volume of a sample used for sorption or porosity measurement etc.)
- Tsang et al. (1996) made the point that, in practice, prediction of areally averaged properties is more useful than attempting to predict local point concentrations.

Sampling optimised

- Sample sizes need to be optimised (on a 'data worth' basis).
- In the same vein, the frequency of data collection needs to optimised (and probably kept constant).
- There needs to be more extensive use of autosamplers and continous datalogging, within cost constraints.
- The argument put by Stegemann et al. (2006) for using smaller samples in wastes is worth consideration, as it may offer considerable advantages over a 'large-bucket' approach.

Constant flow rate

- Flow with pauses is disadvantageous because it introduces periods of hydraulic transients (i.e. non-steady flow) which are not modelled by simple transport models. Therefore a stable and reliable continuous flow should be the aim.
- Change from a constant-head driven flow to pumped flow could be adopted in order to achieve a greater control of flow rate. [79] Having a fixed head system suffers from having to have a (slow) manual adjustment of the outlet level in order to attempt to control the flow. Due to changes in the waste with time (for example from settlement, blinding by bio-films, gas build-up) the K may drop and the flow under a fixed head will also drop. A well-calibrated positive displacement pump should be able to immediately deliver at inlet the flow required and maintain this against a wide range of

[78] See http://www.mines.edu//igwmc/jupiter/ for more information.

[79] Pumps also have drawbacks. The flow rate may still vary due to mechanical reasons or clogging within the pump. Positive displacement pumps are less sensitive to the system hydraulics and can potentially build up very large heads against any blockage so require care, particularly avoiding the possibility of pumping against closed outlet valves. Pumps will of course stop with any power outage unless back-up power is arranged. At the Pitsea site power failure is a distinct issue. However, (albeit with some time delay) the constant head method also relies on power for pumps to fill the upper tanks. Flushing with tap-water from the mains would remove this issue, but for experiments that seek only the transport properties and not an examination of the changes that result from introducing water instead of leachate, it is desirable to flush with leachate.

heads. Although this does not avoid the fact that the system has changed, it is arguably better to keep the key control of flow rate constant and measure the change by means of continuously measuring (logging) the inlet and outlet heads (thereby keeping a record of K against time).

- An interrupted BTC cannot be analysed by the method of moments. The whole experiment takes longer with a pause in (and therefore is more susceptible to the system itself changing). Key measurements taken immediately on recommencement of flow are most susceptible to error. Modelling intermittent flow requires more complex codes that take longer to run (an issue if many code realisations are required, for example when using MC simulations or an automated fitting algorithm).
- A substantial period of constant flow should be established before commencement of any test in order to ensure constant flow at the start of the test.

Achieving a clear input function.

- Achieving a clear input function is relatively easy when the waste is not covered. Otherwise tracer may be introduced to a soil layer in the field, or to a pond or gravel layer in the laboratory. The principle should be set that it is better to have a well-defined input function than to attempt to force a particular function without knowing how well this has been done.

Demonstrating repeatability

- Waste tracer experiments are rarely repeated, but it is important to test the repeatability. This should focus concentration on control.

Measurements at the outlet

- It is worth assessing the value of measuring. There is as yet no firm theoretical basis to deal with flow and concentration data from several outlets. This requires firstly further theoretical development. Secondly an experimental investigation of the outlet condition is needed in order to distinguish between possible dynamical effects at the outlet and differences due to the waste heterogeneity.

Designing experiments for maximum process and parameter identifiability

- Optimised process conditions need to be selected, such as flow rate, and injection time for maximum model and parameter discrimination (e.g. Chapters 5 and 6).
- Methods such as multiple tracers and intermittent flow should be considered for enhanced diagnosis.

- A systematic appraisal is needed for selecting tracers for use in wastes along with examination of their sorption properties.

Minimising input uncertainty

- Ensuring better flow control can be achieved by using pumps.
- Where a pond is used adding tracer by dosing to a mixed tank is a preferable input method.
- Ensuring that the transfer function of the rig is a minor influence in comparison to the waste is important, for example by reducing the volume of fluid stored in the gravel layers to a minimum practical level.

Summary

There are several trade-offs to be made, for example between cost, data-worth, practicality, robustness. These trade-offs must be discussed and modelled in order to make the decisions explicit.

Cost is a critical variable of practical importance. Cost can be split into 'Capital Expenditure' and 'Operating Expenditure'. These will vary greatly depending on the level of equipment already owned by the investigating organisation as well as on the rate-cost for a particular item (bearing in mind that unit rates tend to reduce for larger quantities). As an example of the variability in cost, the relative costs of a single analysis (point on a BTC) to chloride (1) at 2007 prices are roughly: bromide (1), lithium (1) and deuterium (10). In 2007 a single chloride analysis is roughly £2.00 commercially.

There is a need to produce guidance on laboratory experiments for wastes, tracers for wastes and lysimeter-scale experiments for waste, using best practice from soil experiments whenever applicable. A document along the lines of the BGS tracers manual (Ward et al., 1998) would be extremely useful.

12.3 *Hypotheses to test with future work*

The work in this thesis has suggested four hypotheses which need testing:

- Diffusive disequilibrium dominates long-term flushing of many solutes.
- The size and distribution of large, relatively dry unflushed areas is very important for the long-term (and is not characterised by either by laboratory tests or short-term field tests).
- Intermittent flow can cause more of the immobile zone to be accessed (Bendz and Singh, 1999)

- The waste structure dominates the flow-pattern and the transport properties (and implicitly the flow-dynamics are less important).

12.4 *Possible future experiments*

12.4.1 Spatial distribution of structure and properties

A better understanding is required of spatial heterogeneity of properties, of the macro-scale structure of wastes and of the nature, distribution and connectivity of flow paths. Knowledge of the spatial distribution of the flow and transport properties that result is also needed. Attempts have been made to measure the spatial distribution of moisture content in MSW, but these still require refinement (Gawande et al., 2005; Jain et al., 2006). Time domain reflectometry has been used at the laboratory-column scale to visualise the advancing front of moisture in MSWi fly ash (van Praagh et al., 2005) and resistivity (Rosqvist et al., 2005). However, thus far, relatively low resolution imaging has been achieved. In future a range of invasive and non-invasive investigative techniques are required, including for example the use of dyes, geophysics methods and freezing.

A taxonomy of waste types, based on the spatial-structural properties, would be very useful. Types of models and parameters could potentially be related to each group allowing a relatively sophisticated choice of model and parameter. Such a body of work could be used to test the hypothesis that the waste-structure dominates the transport properties, rather than any other effect (for example, self-organised flow structures or leachate-gas interaction).

12.4.2 Analysis of variability (noise) in flow and transport data

There is a need to measure and understand the component contributions to the total noise seen in BTC and flushing data to enable modelling and to seek affordable strategies to minimise it. Time-series data would help in understanding whether there is a high-frequency content (perhaps from bubbles or flow-switching) which can be averaged out by increasing the sample sizes or flow averaging times (depending on how flow is measured).

Increased understanding of temporal flow variability may help to inform different types of models, for example novel channel models that attempt to capture intermittent closure of pathways due to gas blockage.

12.4.3 Analysis of variability (noise) in concentration data

Better knowledge of the contributions to error in concentration data affects the weighting of different points during fitting of models to data.

Instrument error could be established by including in the randomised samples several repeated standards that span the range of BTC concentrations.

If the noise at $C \approx 0$ is also known (far into the tail, or at very early time), then a noise model for the whole system could potentially be interpolated between the two points. This noise model could then be used for a more reliable Monte-Carlo analysis (with an informed estimate of noise, rather than a guess).

A question that needs addressing is how many data points need to be collected in order to obtain a reasonable estimate of the noise (i.e. what is the 'error of the error')?

Standard results for the expectation of the sample variance, $\langle s^2 \rangle$ and the expectation of the variance of the variance, $\langle \mathrm{var}(s^2) \rangle$ of a sample of size N are[80]:

$\langle s^2 \rangle = \frac{N-1}{N}\mu_2$ and $\langle \mathrm{var}(s^2) \rangle = \frac{(N-1)^2}{N^3}\mu_4 - \frac{(N-1)(N-3)\mu_2^2}{N^3}$, where μ_2 is the underlying variance of the noise and μ_4 is the fourth central moment of the noise.

Assuming normally-distributed noise, substitute $\mu_4 = 3\sigma_N^4$ and $\mu_2 = \sigma_N^2$ to give:

$$\langle s^2 \rangle = \frac{N-1}{N}\sigma_N^2 \tag{12.2}$$

and

$$\langle \mathrm{var}(s^2) \rangle = \frac{2(N-1)\sigma_N^4}{N^2} \tag{12.3}$$

The improvement in the ability to estimate the true underlying variance is greatest when the sample size is first increased.

It is also worth giving careful thought to the sampling strategy and its relationship to noise.

12.4.4 Systematic series of experiments

Running a series of tracer tests holding all variables but one constant would potentially reveal a number of important aspects. Varying the length of the waste in the column potentially allows the scaling of dispersion with distance to be established and checked against different process

[80] See for example http://mathworld.wolfram.com/SampleVarianceDistribution.html.

model predictions. Similarly, varying the flow rate allows a check on different process assumptions.

Furthermore, it would be very useful to know how the transport properties of wastes vary with a series of different key variables, for example compression, pore pressure, variable particle size, degradation level, waste type, pH and saturation level. This understanding should both assist conceptual understanding of the system and selection of optimum models. Ultimately this should also indicate practical engineering controls on flushing rates and potentially inform better management or pre-treatment techniques for enhanced flushing.

Examination of the effect of each variable would require the generation of at least several separate BTCs and adequate flushing thereafter, in preparation for another injection. Such a 'campaign' would be best done with automated sample collection and with flow conditions designed such that the test duration is relatively short (otherwise the entire exercise would take an impractically long time). It would be most practical to do this at the laboratory-scale and to sub-sample this and conduct a few tests also at the lysimeter scale.

12.4.5 Emplacement

A layer of tracer-impregnated waste could be placed midway within a column. This would allow a quite different form of tracer investigation than injecting at one end of the column. A better impression of flushing efficiency may be gleaned from this. The mass balance could be much more precisely determined than for the case of flushing an existing contaminant assumed to be uniformly distributed in the waste and leachate at the start of the experiment.

12.4.6 Intermittent flow

The single dataset and model used by Bendz and Singh (1999) inadequately tests the important hypothesis that a greater proportion of the immobile zone is more rapidly flushed when the flow is intermittent. [81] Careful experimentation with intermittent flow and different models which simulate advective as well as diffusive transfer could be used to test this notion more thoroughly.

Whatever the mechanism, it is important to establish whether intermittent flow is more effective at flushing or not. It is worth noting that the results in the literature thus far are inconclusive. There is a discussion of pulsed-pumping by Harvey et al. (1994), in relation to aquifer

[81] Bendz's model needed the size of the 'fast' diffusive transfer zone to be increased when there was transient saturation (hence advective DP transfer) implying that perhaps intermittent flushing allows more of the diffusively flushed immobile zone to be accessed. This is a very interesting and relevant possible mechanism, but there is a need to test how much is due to the model and how much is real. It is therefore given the status of hypothesis here.

decontamination by pump and treat (PAT) using a radial flow model with MIM transfer. They found that if an intermittent regime is compared to a continuous one, for equal volumes extracted the intermittent regime does not remove significantly more contaminant. Indeed, if the stop is too long it removes less. If the pump rates are the same then the total remediation period for the intermittent system will be longer, but will achieve the same removal for less pump time. Rabideau and Miller (1994) also analyzed PAT for a 2D radial system with non-equilibrium transfer (MIM or DP) and linear as well as non-linear (Freundlich or Langmuir) sorption. However, again interruption produced a less pronounced improvement for flushing than might be expected over a timescale of hundreds of months.

12.4.7 Field experiments

Ultimately field scale transport investigation is needed to inform field-scale flushing prediction, even though a lot of insight can be gained by working at smaller scales[82]. This is the most important, but most difficult to achieve data for waste flushing. There is a strong case for tightly controlled and carefully designed long-term field data collection exercises.

On a shorter timeframe, tracer experiments can produce a lot of information about field transport behaviour. Large, relatively dry and unswept dead-zones have been observed and are conjectured to be a common feature. The timescale of diffusive flushing of such zones will potentially be of the order of decades or more and may come to dominate the flushing tail. This is the 'diffusive-dominance hypothesis' and it requires testing.

If there is a significant fraction of waste which has a very large diffusive block time (t_{cb}), sufficiently high resolution tracer BTCs should potentially reveal information about the existence of such a fraction.

One experiment would be to introduce tracer into an unsaturated system by sprinkling. The sprinkling would continue as a 'wash' so that the flow rate is kept continuously constant (rather than left to 'natural' variability governed by the soil moisture balance of the top cover). A further experiment would be to remove the flow control and sprinkle tracer and wash until the tracer is sufficiently established in the waste and then allow the tracer to move at the 'natural' rate. This would potentially be very slow unless a relatively shallow waste was selected.

[82] Experiments in the laboratory can be used to develop techniques, measure processes carefully, give guidance as to what to look for at a greater scale and provide an indication (a 'prior') of what parameter values might be, thereby assisting planning of field or lysimeter scale work. The flushing of laboratory columns may potentially represent an upper bound for what may be flushed at a bigger scale as it is likely that larger heterogeneity features will restrict flushing. However, to what extent laboratory-scale processes and parameters represent the field is an unanswered question that also pervades much of hydrogeology and soil science.

Measurement of precipitation and temperature-humidity would be needed to give estimates of evapotranspiration as well as the total outlet flow of leachate.

The concept of horizontal flushing in saturated wastes could be tested, the motivation for this being that there is anisotropy (horizontal permeability > vertical permeability, i.e. $K_h > K_v$). Horizontal saturated flow may not lead to as many dead-zones as vertical flow, since most major obstructions are conceptualised as compressed layers of plastic in a sub-horizontal orientation.

Horizontal-radial flow properties could be explored by 'echo' tests, which are similar to Aquifer Storage and Recovery (ASR) schemes, whereby a 'bubble' of liquid is injected into an aquifer and then removed again. The parameters established by 'echo' tests could then be used to design more elaborate flushing tests involving either radial flushing well-fields surrounding a central injection well or two opposing lines of wells, with one set pumping and the other abstracting.

12.4.8 Closed-loop systems

Connecting the output of a column to the inlet and running in a closed-loop, rather than running liquid through the column conventionally

The advantages of having a closed-loop system are many. Such a test would be self-contained and would not require disposal of any effluent during the test. Any dynamic effects due to water plus tracer suddenly replacing leachate are removed. The cell should therefore remain close to the state (pH, Eh, TDS) in which the leachate has found equilibrium with the waste mass.

This method may be useful for measuring transport through clay liners for which it is not obvious how much of the porosity is involved in transport (some of which may be irreducible). It is worth investigating whether there is exclusion of different solutes due to the solute molecular size or by electrostatic repulsion (e.g. anion exclusion).

Two further experimental problems for BTC data from waste column tests are potentially overcome. Firstly, the waste often compacts down so that the length to diameter ratio of the column is often rather small, somewhat stretching assumptions of 1D flow. Secondly, there is often a large volume of liquid in the basal gravel relative to the volume of liquid in the waste. It might be that the experimental apparatus exerts a considerable influence on the results. The results for a rapidly recirculating closed-loop system should (after some time) become dominated by diffusive effects (if the diffusion-dominated hypothesis is accepted) and the total

porosity. The parameters describing diffusion times and porosities, which are the most desirable properties to know with regard to the long-term behaviour, can be recovered this way. AD-DP processes could be modelled in a closed loop system using a modified version of DPD-PULSE (in which the solute from the last cell is passed back to the first cell).

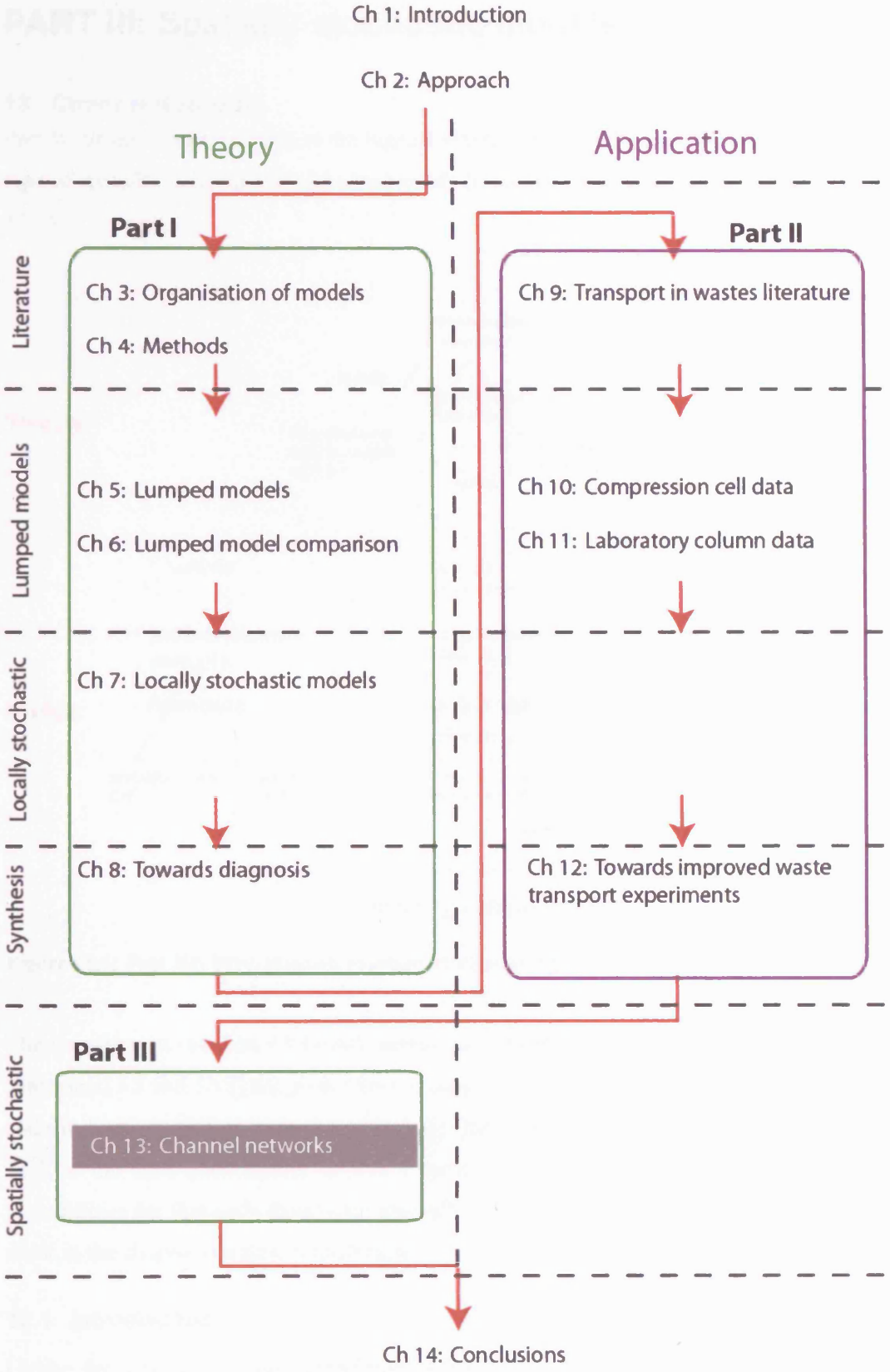
12.5 *Conclusions*

Increased understanding of transport processes in wastes will come from an iterative process of experiments, modelling and suggested enhancements and cycling over further experiments. Key enhancements to existing experiments have been suggested, from the most basic to major adaptations. Possible future experiments beyond these existing experiments have also been discussed.

The role of modelling will be integral to this experimental refinement and uncovering of processes and parameters. Modelling must delve into the 'black box' of process behaviour only as far as is needed to make reliable predictions. Simple physically-based process models are arguably at the right level to do this. This requires both a set of hypothetical field and management scenarios to test and the creation of a new toolkit for making risked predictions.

The three most important issues for modelling of wastes are:

- A reassessment of the value of historical data and an analysis of to what extent many existing flushing problems can be addressed through *interpolation* from existing and future long-term datasets (and holding on to the measurement of both the mean prediction and the uncertainty in the mean). This is arguably greatly more desirable than interpolating by using simple models to fit data, for example the exponential models used by Reitzel et al. (1992).
- Establishing key simplifications in order to reduce the complexity of poorly constrained boundary conditions and multiple interacting processes. This will need to be based on well-justified physical grounds, perhaps through intentional suppression of processes during experiments.
- Moving forwards to provide more thorough predictions in a manner that carefully and adaptively assesses the predictive uncertainty. Such a method could be based on multiple working hypotheses ('pools' of simple process models such as those developed in Chapters 5 and 7) the confidence in each of which can be continuously refined (in either a formal or pseudo Bayesian manner in confrontation with data and taking into account model complexity).


Ch 1: Introduction
Ch 2: Approach
Theory
Application
Part I
Part II
Literature
Ch 3: Organisation of models
Ch 4: Methods
Ch 9: Transport in wastes literature
Lumped models
Ch 5: Lumped models
Ch 6: Lumped model comparison
Ch 10: Compression cell data
Ch 11: Laboratory column data
Locally stochastic
Ch 7: Locally stochastic models
Synthesis
Ch 8: Towards diagnosis
Ch 12: Towards improved waste transport experiments
Part III
Spatially stochastic
Ch 13: Channel networks
Ch 14: Conclusions

PART III: Spatially stochastic models

13 Channel Networks

Part III of this thesis operates at the highest heterogeneity scale, specifically by analysis of one type of spatially stochastic model which needs further development – a channel network (Figure 13.1).

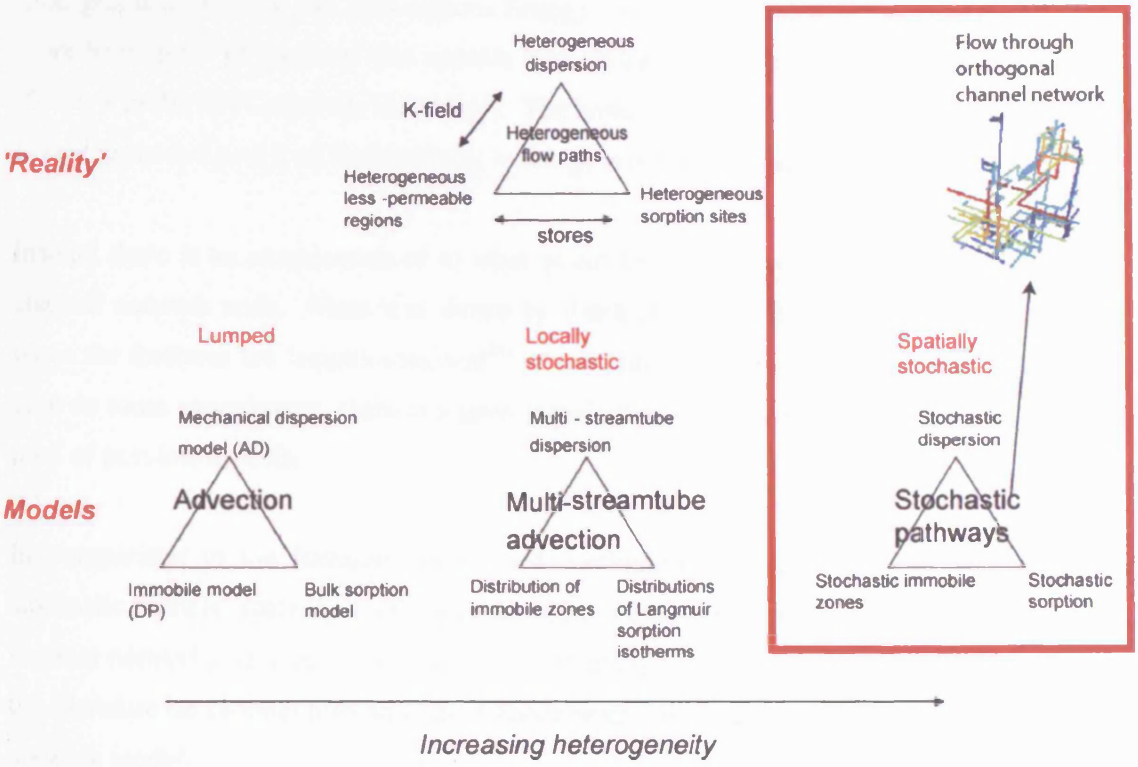


Figure 13.1: Part III: Flow through spatially stochastic system modelled by channel network.

The introduction (Section 13.1) and section on channel network models and percolation theory (Sections 13.2 and 13.3) are predominantly review material. The introduction to the HyperConv code in Section 13.5 is reviewing this code (the code was written outside this thesis). Table 13.3 is the first contribution in this Chapter; a more comprehensive set of characteristic probabilities for this code than were originally devised by its author. Thereafter, the remaining work in the chapter is a new contribution.

13.1 *Introduction*

Unlike the transport models considered in Parts I and II of this thesis, it is the structure and nature of the flow system which is analysed in this chapter. This is because in order to understand why channel networks behave differently to continuum systems requires closer inspection of the most basic behaviour of a network. It is this fundamental aspect that is the focus here. Once the network is understood, it is relatively straightforward to describe how bulk

and apparent local permeability of channel networks relate to the underlying system. A further contribution from this author to discussion of implications for permeability is contained in Black et al., 2006[83]. Analysis of the transport behaviour, building upon the work here would make a useful future contribution.

Models are not confronted with data here, nor is it appropriate to apply the three methods of analysis that are applied to the lower levels of heterogeneity (i.e. the method of moments, late-time gradient analysis and least-squares fitting). This is because this model first requires the more basic level of appraisal that appears here (in particular the basic analyses of connectivity (Level 1 in the 'BTC analysis hierarchy'). The basic statistics of the network are examined (this is analogous to Level 2 of the hierarchy, although is not actually analysis of a BTC).

Instead there is an examination of to what extent key phenomena may emerge from a simple channel network code. Since it is shown by Black et al. (2006) that discrete fracture models, when the fractures are 'equidimensional[84]' or continuum models struggle to simulate the effects seen in some experiments, there is a good justification for examination of a new entrant to the pool of possible models.

In comparison to the literature on lumped continuum flow and transport models, locally-stochastic models, spatially continuous stochastic permeability and discrete fracture networks, channel networks have received relatively little attention. Therefore a contribution is made to the literature on channel networks, by detailed examination of the behaviour of a new channel-network model.

The original driver for this work was flow and tracer observations made around underground research laboratories (URLs), which are being examined for possible use as sub-surface nuclear waste repositories. Tsang and Neretnieks (1998) summarised flow and tracer experiments that can be adequately (or only) explained by channel-based conceptual models. John Black (personal communication) builds a hydrogeological case for a need to conceptualise a fractured crystalline rock environment in terms of sparse, elongate channels. Some of the evidence cited by Black for this conceptualisation as an alternative to EPM or fracture-network models is given in Table 13.1. The primary phenomena that are suggested to be potentially incompatible with conceptualising the system as an equivalent porous medium are:

(i) Localised inflow points to underground openings (rather than planes or areal seeps).

[83] This report was for Svensk Kärnbränslehantering AB (SKB) and should eventually become an open-source 'R-series' report.

[84] 'Equidimensional' means an aspect ratio of 1 (e.g. circular or square instead of elliptical or rectangular). Fracman assumes equidimensional fractures, for example.

(ii) Development of apparent 'skins' around underground openings when interpreted by continuum models.

(iii) Apparent 'compartmentalisation' of heads (rather than a smooth predictable variation).

Channel-network models might be able to represent all three phenomena. Localised tunnel inflow is an inevitable result of assuming channelled flow. The second observation requires more investigation. The existence of a 'skin' by its name infers a zone of enhanced or reduced transmissivity around the opening (top of Figure 13.2). A regression of log(radius) and head will predict negative or positive heads at the inner boundary. On one hand, enhanced permeability might occur due to an Excavation Damage Zone (EDZ) and this would predict a negative skin. However typical interpretations give large *positive* skins (Gale et al., 1982), perhaps inferring a zone of reduced permeability, contrary to what might be expected following blasting, stress relief and any other construction disturbance. In practise, neither the thickness nor the transmissivity of such a zone has ever been established with any confidence. Alternatively, skin may result due to different distributions of head (typical modelled distributions are illustrated at the bottom of Figure 13.2).

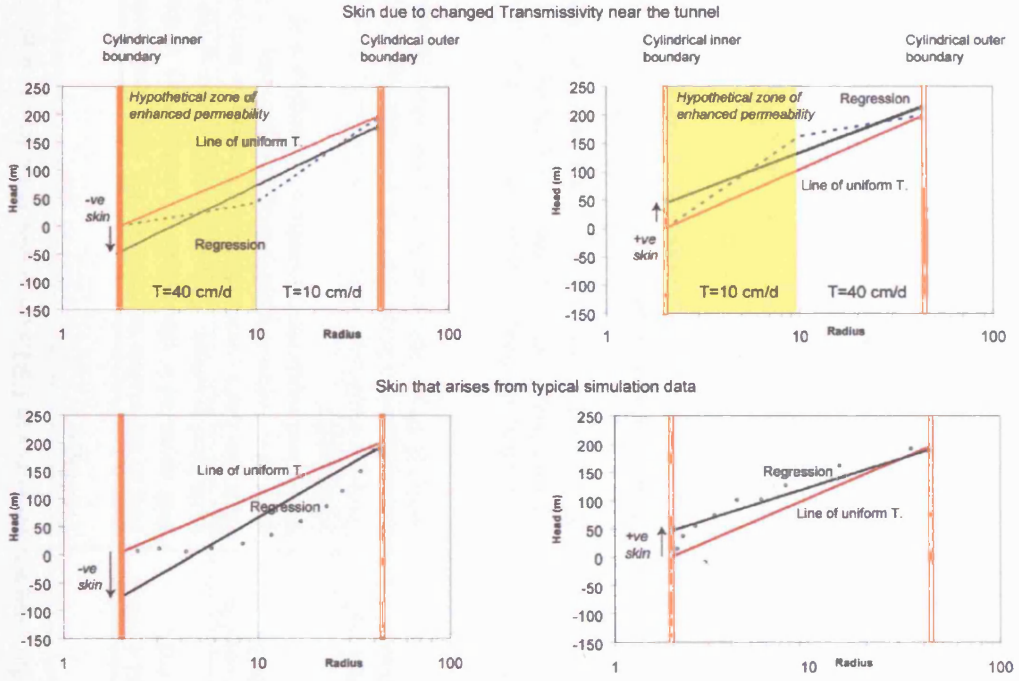


Figure 13.2: The effect on the head-radius regression of a band of enhanced or reduced transmissivity in ideal cylindrically symmetric flow and of typical channel-network simulation head data (as seen in Section 13.8.2). Note dashed line gives the head at different radii.

Table 13.1: Summary of evidence for channel networks in URLs (evidence from Black et al. (2005) and Black et al. (2006)).

Experiment (location)	Author	Findings
3D migration experiment (Stripa mine)	Birgersson et al. (1994) cited in Black et al. (200)	Few of the visible fractures on the mine drift surface participated in flow. Tracers injected at one point in the drift appeared in places tens of metres away in the drift in a manner that was seemingly unrelated to the large-scale head gradient. This apparently points to the existence of drift-parallel channels.
Simulated drift experiment (Stripa mine)	Olsson (1994) cited in Black et al. (2005)	Flow reached a central borehole which was surrounded by five boreholes with substantially lowered heads. This must have required a fairly long channel with <1.7 m width to pass between the outer boreholes. The entire region was excavated to create a 'validation drift'. Flows to this drift were very localised (6 inflow areas in a 50m length).
Large-scale macro-permeability experiment (Stripa mine)	Gale et al. (1982) and Witherspoon et al. (1980)	During 'ventilation' abstraction of moisture from the drift, packer testing in ten boreholes showed a consistent head drop towards the drift, but also with some large departures from the general radial trend. A large 'skin' (order of 40 m) needed to be invoked when interpreting the data using a continuum flow model.
Grimsel Shear Zone pumping and tracer tests (Grimsel)	Meier et al. (2001)	Geostatistic inversion of cross-hole flow tests appeared to give channel-like distributions of transmissivity (40 different fields were estimated that all honoured the flow and head data, but gave different transport predictions although had a coarsely similar channelling pattern).
USGS Fractured Rock Hydrology Research Site (Mirror Lake)	Day-Lewis (2000)	Simulated annealing was used to construct a 3D K distribution, which could be interpreted as having a (sinuous) high-K path (which could perhaps be called a channel).
Kamaishi mine (Japan)	Sawada et al. (2000)	Compartments ('zones') of the order of 10 to 40 m in scale could be defined on the basis of similar head and rapid transmission of hydraulic signals within compartments. The compartments showed no obvious relationship with the organisation of fractures.
Sellafield exploratory boreholes (UK)	Black et al. (2006)	Drilling and multilevel head measurements apparently showed compartments of different heads at different depths.
TRUE Block Scale experiment (Äspö)	Poteri et al. (2002) Rhén et al. (1997)	Although heads have been interpreted due to 'structures' in a hydrostructural model, in order to reproduce observations more and more flow barriers have to be invoked. The data might be better interpreted by conceptualising compartments.

Assuming steady-state flow, there are a number of analytical solutions for flow to linear features with cylindrical cross-sections. Goodman et al. (1965) gave an analytical solution that takes into account the proximity of the water table to a tunnel. At the very simplest level radially symmetric flow can be assumed and flow modelled using a modified Thiem equation (applicable to both deep tunnels and wells).

$$h = h_0 + h_s + \frac{Q}{2\pi LK}\log(r/r_0) \tag{13.1}$$

where h_0 is the head at the tunnel, h_s is 'skin' and L is the length of the tunnel. A regression of h against $\log(r/r_0)$ will give a slope, α and the intercept of the regression at $r = r_0$, which will differ from h_0 by h_s. $h = (h_0 + h_s) + \alpha\log(r/r_0)$.

Barker (1988) generalised this showing that flow can be interpreted in terms of different assumptions about the dimensionality of a continuum. Flow might occur through a linear feature, such as a single channel and therefore have a flow dimension, D_F, of one. Flow might be confined to a layer and converge radially (i.e. D_F=2) or flow might converge spherically around the abstraction point (D_F=3). Barker (1988) realised that there might exist intermediate dimensions of flow between these dimensions, in other words he proposed use of *non-integer* flow dimensions. At the simplest level, the area enclosing a volume at a given radius is given by $A = Cr^{D_F-1}$, where C is a constant.

It is possible, given a distribution of $h(r)$, to calculate D_F at any point. Equation 40 in Barker (1988) can be logged, differentiated and rearranged to yield:

$$D_F(r) = 2 - \frac{d\ln h}{d\ln r}\left[1 - \left(\frac{r_0}{r}\right)^{2-D}\right] \tag{13.2}$$

Caution needs to be applied using this meaningfully, however, since dimensionality is here interpreted relative to distance from a point with inward flow. The flow system might behave symmetrically between two F2F boundaries, yet any symmetric perturbation away from linear flow will change the dimensionality in opposite senses for the inner and outer boundaries.

Table 13.2 describes how nodal frequency and head would be expected to develop for a continuum system.

Table 13.2: Prediction of nodal frequency and head assuming continuum system (after Barker (1988:Equation 40)).

	F2F (D_F=1)	**Cylindrical** (D_F=2)	**Spherical** (D_F=3)
Nodal frequency	$\propto r^0$ i.e. constant	$\propto r$ i.e. linear	$\propto r^2$ i.e. log-log linear
Head	$h = h_0 + C_1(r - r_0)$ i.e. linear	$h = h_0 + C_2 \ln(r / r_0)$ i.e. log-linear	$h = h_0 + C_3\left(\frac{1}{r} - \frac{1}{r_0}\right)$ i.e. linear in $\frac{1}{r}$

where $C_1 = \frac{h_1 - h_0}{r_1 - r_0}$, $C_2 = \frac{h_1 - h_0}{\ln(r_1 / r_0)}$ and $C_3 = \frac{h_1 - h_0}{1/r_1 - 1/r_0}$ are constants and subscript 1 denotes outer boundary and 0 denotes the inner boundary.

Because of the difficulty in interpreting plots of D_F , the linearity or log-linearity of the nodal frequency and head in channel networks are tested according to Table 13.2 as well as the simplest option of fitting with continuum head or nodal frequency models. The advantage of the simplest option is that both the boundary heads as well as the dimensionality can be altered. This might not be immediately obvious as an advantage, but it will be shown in Section 13.8.2 that there are occasions where very steep head gradients characterise some boundaries yet the remaining system conforms to a given flow dimension.

13.2 *Channel network models*

There has been some previous work on channel networks, but it is a relatively small literature. Gylling et al. (1999) described the channel network model CHAN3D, which is a flow and transport code. The network is restricted to 6 sub-channel connections per node (as is the case for a cubic grid). Sub-channels are allocated non-correlated, lognormally distributed conductivities that combine with 'length' variations to give produce a distribution of conductances. In practise there are therefore only conductance variations and no variations in the true length of a subchannel (except for changing the grid size). Gylling et al. (1998) used CHAN3D to simulate a pumping test at Äspö HRL. Moreno et al. (1997) constrained the network to a cubic grid to model tracer tests at Äspö. Selroos et al. (2002) compared CHAN3D with a stochastic continuum and a DFN model.

The way in which fracture network systems might 'collapse' to channel networks is a moot point. Fracman (Dershowitz et al., 1995) converts networks of equidimensional fractures into channels that connect the midpoints of the intersections. Cacas et al. (1990) applied such a

concept to modelling flow near a mine in granite (at Fanay-Augères). Another leading code that does this is ConnectFlow, which is based on the equidimensional DFN model, NAPSAC (Hartley, 1998) and the porous media model, NAMMU (Morris and Watson, 1999). None of these models produces networks of elongate and sparse channels since they are based on equidimensional fractures. It may be the case that multiple channels occur within each fracture plane and it is not appropriate to idealise each fracture with a single channel.

A cubic channel network can potentially perform in the same way as a stochastic continuum model (Moreno and Tsang, 1994), but with K being allocated on a cell rather than sub-channel basis.

There is therefore a gap in the literature for analysis of spatially correlated conductors. Moreover, intersection probability of longer fractures is increased, so percolation may occur for less dense (sparser) systems than might have been conceptualised for 'classical' channel networks or for equidimensional DFNs (Black et al., 2006). This gap is partially addressed here, for a specific network.

One further piece of theory is required in order to understand these systems – percolation theory.

13.3 *Percolation theory*

Percolation theory has a broad range of physical applications (Sahimi, 1994), and has been applied to subsurface flow and transport problems (Berkowitz and Balberg, 1993). Percolation networks are generally analysed by randomly populating a plane (for 2D) or volume (for 3D) with conducting 'bonds' (sub-channels). These bonds might be constrained to a certain lattice of sites (nodes), having z neighbouring nodes which could potentially be connected (z is the 'co-ordination number'). In a 3D cuboidal lattice, z =6. Alternatively there might be no lattice confinement and bonds may be placed in space with random position and orientation. This is called 'continuum percolation' (Hunt, 2005).

The percolation threshold, P_C, is established by increasing the bond occupancy, P, (or density for a continuum system) by randomly generating bonds between sites until two of the domain boundaries are connected (traditionally established between planar boundaries). Ideally P_C is established for infinitely large systems, but for all but the simplest systems it must be established numerically within a finite domain. For finite systems in which the domain size is a small multiple of the bond length, the percolation threshold is spread over a range of P.

A major preoccupation has been with the scaling of K and other network properties with P, and P_C near to the threshold (i.e. $K \propto (P - P_C)^{\kappa}$, where κ is a constant that depends on the dimensionality of the system).

Margolin et al. (1998) examined the effect on flow and transport of having a stochastically generated channel conductance in a classical network. Berkowitz and Braester (1991) examined transport by using a random-walk transport code for transport through variable aperture near-critical percolation fracture networks. There is literature on the percolation behaviour of different shaped fracture networks (e.g. Robinson (1984) for equidimensional square fractures in 3D).

Where there is no spatial correlation between bonds, the system is here called a 'classical percolation network'. This is to distinguish it from networks with some directional bond correlation (potentially producing a range of channel sizes). There has been relatively little work on non-classical percolation. De Dreuzy et al. (2000) considered percolation through systems with randomly generated 3D ellipsoids with power-law size distributions, in which the eccentricity could be adjusted. Garboczi et al. (1995) evaluated same-size ellipsoids for different aspect ratio. The effect of variable object sizes was to spread P_C, as was having a finite domain. No studies of grid-confined percolation for elongated channels were found, so the work here is new.

13.4 *Approach*

For this study a variable-length channel network model called 'Hyperconv' is examined (Black et al., 2005)[85]. The C++ code was originally written by Dr Peter Robinson of Quintessa Ltd, but this author has added spherical boundary conditions and further output statistics (see Appendix 1, page 442, for more details). HyperConv was written in order to examine the effect of sparse channel networks in the presence of 'long' channels on flow to subsurface repositories and the consequent evaluation of permeability and apparent 'skin'. The work here develops the initial scoping work of Black et al. (2005) to a considerably more detailed level.

Two central hypotheses about channel networks are examined here:

[85] The code was called HyperConv because it was originally written in order to search for supposed 'hyperconvergence' of flow towards underground openings, in the manner of Channel hypothesis 1.

Channel hypothesis 1: Sparse channel networks might 'grow' away from single points of connection with an opening in a spherical manner until sufficient interaction has occurred with the growing 'bubbles' from other points of intersection and a more cylindrical flow dimension is experienced further out. Such a network growth scenario might account for the apparent positive skins observed in some URL experiments.

Channel hypothesis 2: Sparse-channel networks may organise in such a way as to cause compartmentalisation of head, with compartments being created between certain key channel linkages.

In order to examine these hypotheses a systematic examination is performed. Firstly, the nature of the code for network generation is examined. Secondly it is shown that, due to the stochastic nature of the networks, it is necessary to generate large ensembles of networks. Thirdly, in order to test the network growth behaviour of '*Channel hypothesis 1*', three different boundary conditions are considered in turn:

- Flow between an inner cylinder (tunnel or well) to an outer cylindrical boundary - 'cylindrical' flow.
- Flow between two opposite faces on a cube (perhaps relevant to long distance flow, laboratory columns or between fracture zones) - 'face to face' (F2F) flow.
- Flow between an inner sphere (perhaps a packer or a single inflow point) to an outer spherical boundary - termed 'spherical' flow.

The cylindrical flow case is examined first, as it is more commonly experienced. The prediction of skin is used as a rough classification of network behaviours. This classification is then used to fix a set of network generator parameters that are then applied to the other boundary conditions, in order to investigate network growth with consistent and relevant parameters.

The behaviour of the network is examined by a novel (yet simple) method. The ensemble frequency of nodes which are on the percolating cluster and of nodes which have flowing channels connected to them are calculated as histograms in space. This allows a clearer representation of how the network is developing away from the various boundaries than is achieved by eye when examining visual images of each realisation. Distributions of different properties have been reported as flow rates in different sub-channels (Margolin et al., 1998) but apparently without attempting to represent the spatial distribution of such properties within the network.

Finally, a FORTRAN code is used to examine compartmentalisation of single realisations of networks, in order to examine channel hypothesis 2.

13.5 *Facets of the Hyperconv code*

13.5.1 Network generator characteristics

HyperConv produces a channel-network in which channels are constrained to an orthogonal equi-dimensional grid. A 60 × 60 × 60 grid is adopted here. The spacing between adjacent nodes is given a dimension of 1.5 m. This constraint was to a certain extent a modelling convenience. Channels are created by applying a simple generator algorithm. Figure 13.3 defines the terminology adopted here by showing short lines of adjacent nodes in one direction.

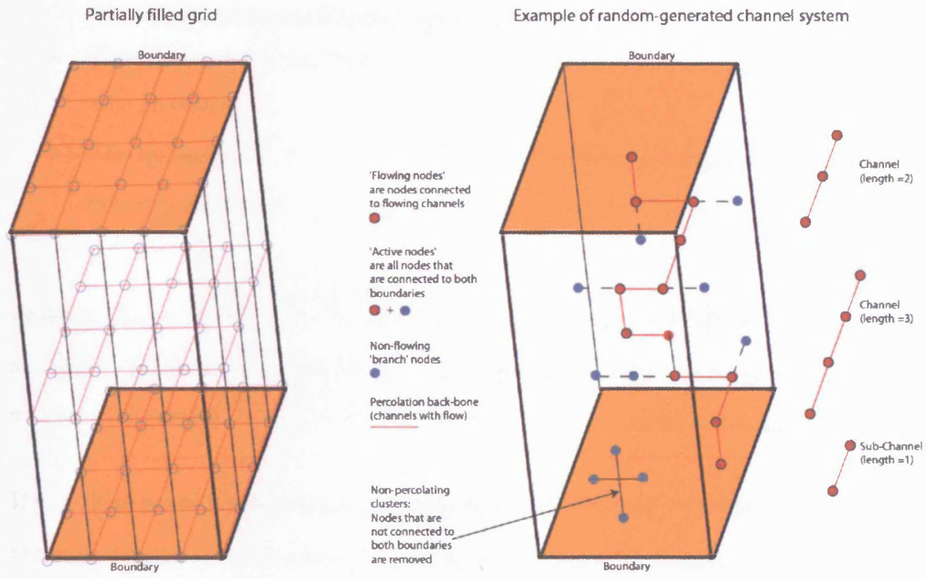


Figure 13.3: Schematic diagram giving the HyperConv terminology (left diagram indicates the grid-structure of the nodes and right diagram indicates a typical network)

'Channels' are created by having sequential bonds ('sub-channels') between sites ('nodes'). A channel is here defined as a sequence of sub-channels in a single direction. When channels in different directions intersect and connect these are referred to as 'pathways' or the 'connected network'.

In 'classical percolation', the probability of a sub-channel existing, P_{ON}, is the same between every pair of nodes. There is therefore no correlation structure and the final network can be populated with many, potentially unrealistic, short channels. The generator used here considers whether to create a sub-channel between each pair of nodes along a line in turn. The probability of there being a sub-channel between two nodes is increased if there was a sub-channel between the previous two nodes. In this way some correlation is introduced along each line of nodes. Specifically, the generator follows the logic that is repeated for each line of nodes and for all three directions:

- The first (boundary) sub-channel is generated (between nodes 1 and 2) with a probability, P_{ON}. This ensures that all sub-channels have the same probability of being generated (without there being any effect due to the boundaries)
- The next sub-channel (between nodes 2 and 3) is generated with probability P_A if there was a first sub-channel between nodes 1 and 2.
- If there was no first sub-channel, then a sub-channel between nodes 2 and 3 is generated with probability P_N.
- The application of P_N if there was no previous sub-channel and P_A, if there was a previous sub-channel, is repeated sequentially across all nodes in the line.

Because, $P_{ON} = \overbrace{P_{ON}P_A}^{\text{Previous=ON}} + \overbrace{(1-P_{ON})P_N}^{\text{Previous=OFF}}$, rearranging gives $P_{ON} = P_N/(1-P_A+P_N)$ (Black et al., 2005). This is used for generating the first sub-channel. P_{ON} can be interpreted as a 1D measure of channel density (i.e. number of sub-channels / total number of possible bonds).

The resulting geometric distribution of channel lengths for different P_A is shown in Figure 13.4. The mean lengths of the channels ($\overline{L}$), the mean gaps between channels ($\overline{G}$) and the modal channel length as functions of P_A and P_N are shown in Table 13.3. Also shown in Table 13.3 is the probability of a given sub-channel belonging to a channel of length N. This is plotted in Figure 13.5 for different P_A. This might be a useful measure as it will relate to issues such as what length and connectivity of channel a packer might be sampling from, or how likely a drift is to intersect a given channel.

While for coding purposes the generator is readily characterised by these two (independent) probabilities P_A and P_N, the characterization of the resulting fracture distributions is potentially better characterised in alternative ways. Table 13.4 illustrates how it would be possible to parameterise the system differently. Instead of a choice of two out of P_A, P_N and P_{ON}, two out of $\overline{L}$, $\overline{G}$ and P_{ON} are chosen. This offers the chance to use the more easily measured variables. For example, $\overline{L}$ is related to fracture trace lengths and P_{ON} to linear fracture density.

The range of possible generator outcomes is given in Table 13.5. A 'classical' percolation network can be generated by setting $P_{ON} = P_N = P_A$.

Table 13.3: Distributions of channel lengths and gaps.

Distribution of	Probability Density Function	Mean	Variance	Mode	Comment
Lengths of channels, L	$P\{L\} = P_A^{L-1}(1-P_A)$	$\bar{L} = 1/(1-P_A)$	$\sigma_L^2 = P_A/(1-P_A)^2$	1	Geometric distribution
Lengths of gaps, G	$P\{G\} = P_N(1-P_N)^{G-1}$	$\bar{G} = 1/P_N$	$\sigma_G^2 = (1-P_N)/P_N^2$	1	Geometric distribution
Lengths of channels containing a given sub-channel, C	$P\{C\} = CP_A^{C-1}(1-P_A)^2$	$\bar{C} = (1+P_A)/(1-P_A)$	$\sigma_C^2 = 2P_A/(1-P_A)^2$	$-1/\ln(P_A)$	

Table 13.4: Relationship between probability generators and mean lengths (and P_{ON}).

	Probability-based characterisation		Length-based characterisation	
Parameters	P_A	P_N	$\bar{L}$	$\bar{G}$
Relation to alternatives	$\bar{L} = \frac{1}{1-P_A}$	$\bar{G} = \frac{1}{P_N}$	$P_A = 1 - \frac{1}{\bar{L}}$	$P_N = \frac{1}{\bar{G}}$
Probability of a sub channel	$P_{ON} = \frac{P_N}{1-P_A+P_N}$		$P_{ON} = \frac{\bar{L}}{\bar{G}+\bar{L}}$	

Table 13.5: Limits to derived parameters (with generator parameters P_N and P_A in parentheses).

Parameter	Maximum	Minimum
$\bar{G}$	∞ (P_N=0)	1 (P_N=1)
$\bar{L}$	∞ (P_A=1)	1 (P_A=0)
P_{ON}	1 (P_N=1, P_A=1)	0 (P_N=0, P_A=0)

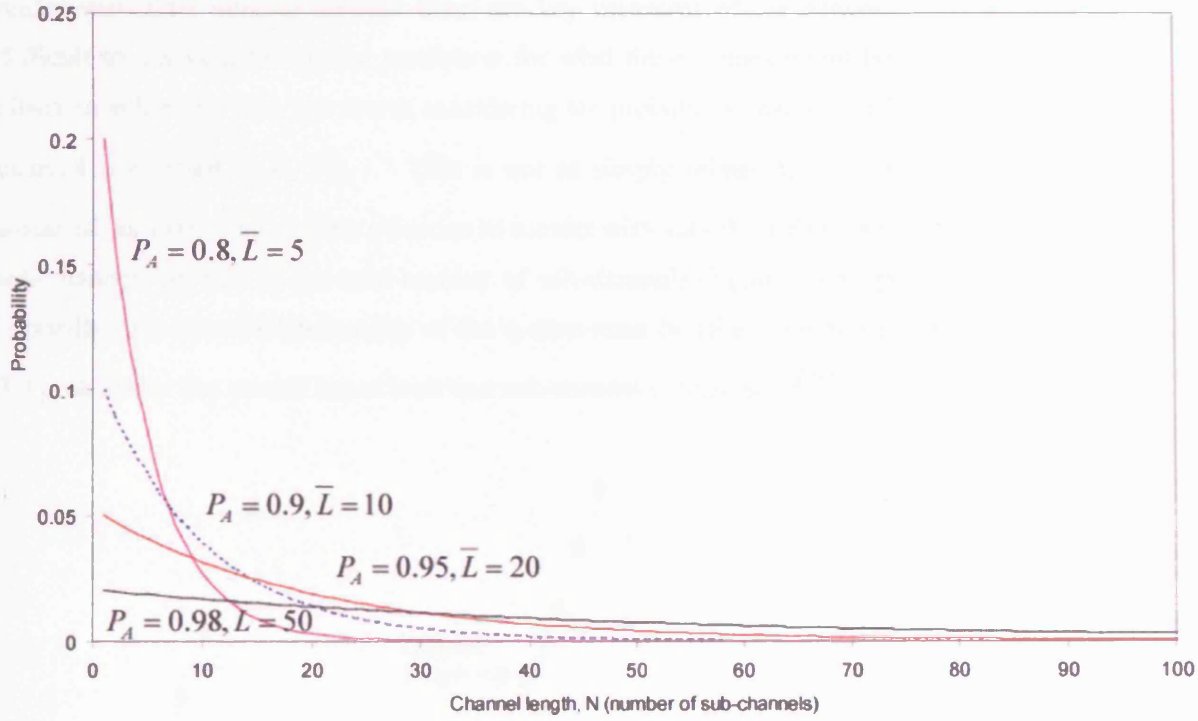


Figure 13.4: Probability distribution of channels of length N, for different values of P_A or average length, $\bar{L}$.

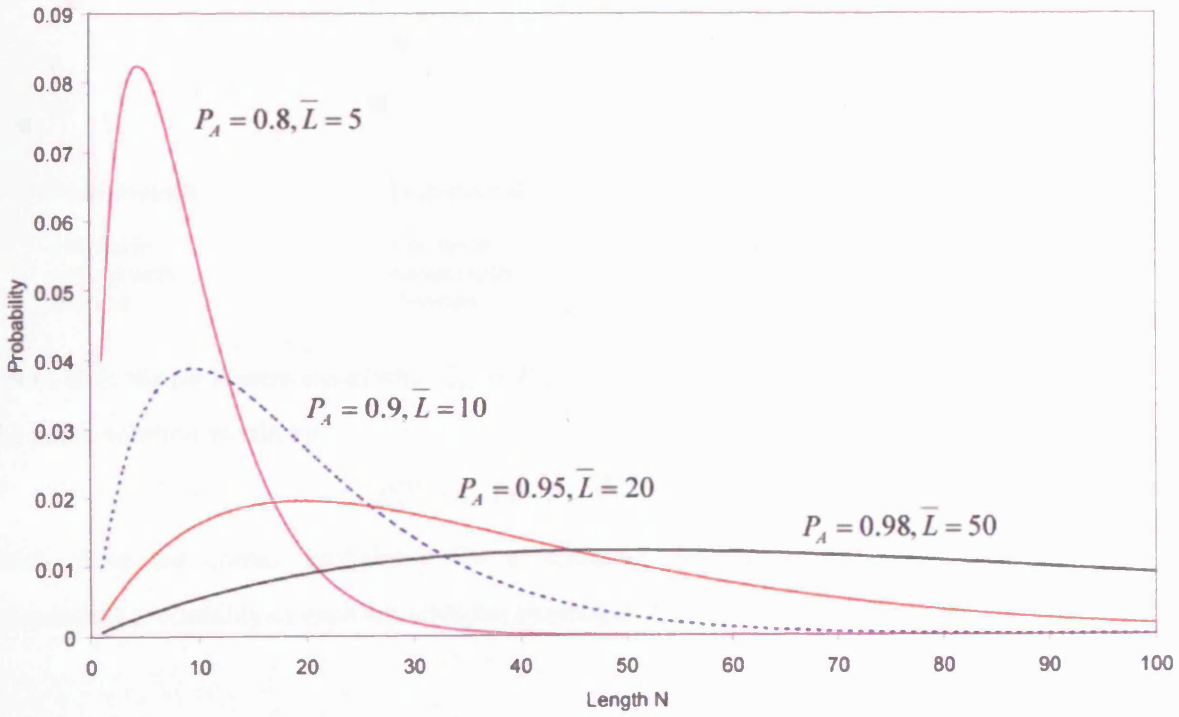


Figure 13.5: Probability that a sub-channel belongs to channel of length N (lengths in number of sub-channels), for different values of P_A or average length, $\bar{L}$.

At any given point, the frequency of nodes connected to both boundaries and the frequency of nodes with flow passing through them are key measures of the network.[86] It is, however, difficult to derive a theoretical prediction for what these values might be. In order to move closer to achieving this, it is worth considering the probability that a node has at least one sub-channel connected to it, $\hat{P}_{ON}$. This is not as simply related to P_{ON} as might be initially assumed, as firstly the number of nodes in contact with sub-channels depends on the spacing of sub-channels as well as the total number of sub-channels (Figure 13.6 illustrates this simply). Secondly, the three-dimensionality of the system must be taken into account in order to give a 3D probability that a node has at least one sub-channel connected, $\hat{P}_{ON}^{3D}$.

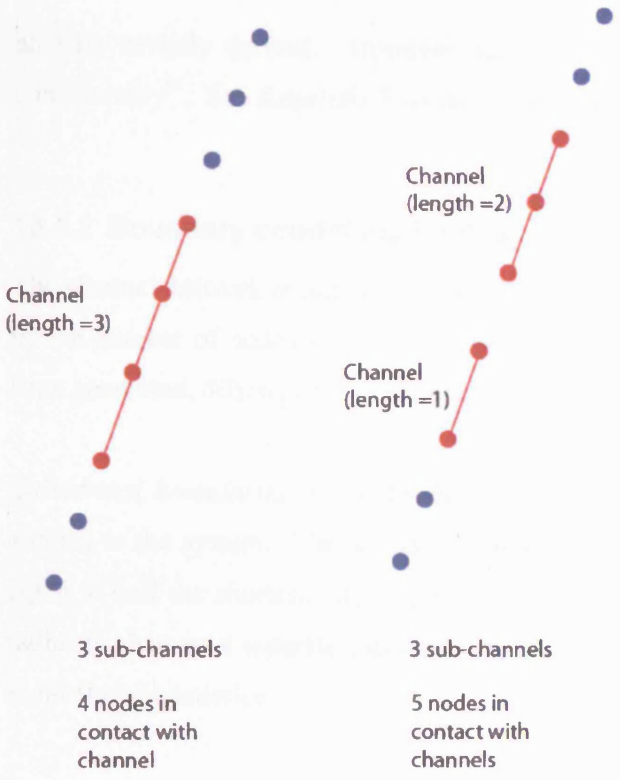


Figure 13.6: Simple illustration of why $\hat{P}_{ON} \neq P_{ON}$.

The naïve solution would be

$$\hat{P}_{ON}^{3D} = 1-\left(1-P_{ON}\right)^6 \tag{13.3}$$

which gives the correct probability for a 'classical' percolation network, for which the independent probability of each sub-channel existing is P_{ON}.

[86] Notice that flowing nodes are the sub-set of active nodes which have finite flow in them. The remaining nodes (active – flowing) are therefore connected to both boundaries but are in some way dead-end branches from the main percolating cluster.

However, this does not take into account the correlation of the channels along each line that is created by this generator. Calculation of $\hat{P}_{ON}^{1D}$ therefore cannot be done at a single point, but requires consideration of channel clustering.

The solution to this problem for nodes and sub-channels along a single line is:

$$\hat{P}_{ON}^{1D} = P_{ON}(2 - P_A) \tag{13.4}$$

Taking into account the three-dimensionality of the problem, gives:

$$\hat{P}_{ON}^{3D} = \hat{P}_{ON}^{1D} \left(3 - 3\,\hat{P}_{ON}^{1D} + \hat{P}_{ON}^{1D}\right) \tag{13.5}$$

As a check, for 'classical percolation', substituting $P_A = P_{ON}$ gives $1-(1-P_{ON})^6$, which can also be trivially derived. However, for the non-classical generator, the solution can differ considerably[87]. See Appendix 3 for derivations of these probabilities.

13.5.2 Boundary conditions for flow

The channel network is initially generated in a box for which the length of each side is defined by the number of nodes and the distance between each node for each direction. Once this has been generated, different boundary conditions can be applied, namely:

Cylindrical boundaries: the outer boundary is defined by a 'cylindrical cookie-cutter' being applied to the system. The boundary exists on a radius centred about the x-axis and of length equal to half the shortest side length. The inner boundary is created in a similar way, but the radius is given as a separate specified variable. The two remaining side boundaries are defined as no-flow boundaries.

Face to Face (F2F) boundaries: the system remains box-shaped and the four 'side' boundaries are defined as no-flow boundaries.

Spherical boundaries: both the outer and inner boundaries are defined by a 'spherical cookie-cutter' being applied to the system. There are no side boundaries.

For the cases with side boundaries it would be a useful future exercise to code a periodic side boundary in order to assess its effect.

---

[87] The expression for $\hat{P}_{ON}^{3D}$ has been checked numerically in Visual Basic by simulating a 3D network and counting the nodes in contact with sub-channels.

13.5.3 Conductance

Channel conductances, C_S, are selected randomly from a distribution as follows: $C_S = C_B \times 10^{us}$, where $C_B = K_{\text{target}} \times \Delta L^2 / P_{ON}$, u is a randomly generated unit normal variable, s is the SD of the $\log_{10}(C_s)$ and K_{target} is the bulk conductivity that should result from a F2F simulation when s=0.[88] Note that the conductances are randomly allocated and are uncorrelated with channel length.

By setting the SD of $\log_{10}(C_S)$ to zero (s=0), a constant conductance network can be produced. In this way it is possible to compare and therefore distinguish the significance of the network structure and the superposed effect of varying channel conductance. For all the simulations presented here, each individual channel is given the same conductance as the first sub-channel that is generated in the line (i.e. conductivity does not vary along a channel).

13.5.4 Network statistics

The HyperConv code calculates 'skin', using the regression based on Equation (13.1) in two ways – from heads at flowing nodes and from heads in the larger set of active nodes.

Three types of nodes are defined in Figure 13.3.

- **Active nodes** – nodes which are connected to both upstream and downstream head boundaries (but are not on the boundaries). Number of active nodes is denoted by N_A.
- **Flowing nodes** – nodes which have at least one sub-channel connected to them with a head drop greater than 1.10^{-8} m (but are not on the boundaries).[89] Number of flowing is denoted by N_F.
- **Branch nodes** – these are active nodes with no flow (i.e. $N_B = N_A - N_F$).

[88] Gelhar (1993) gave the F2F result for a 3D system with a lognormal distribution of K as $K_{eff} = K_g[1 + s^2/6]$, where K_g is the geometric mean. A verification of HyperConv was achieved by modifying the generator to calculate $C_S = C_B \times e^{us}$ and fully occupying the network and testing the computed K. For $s = 0$ Hyperconv gave K_{F2F}=2.00×10^{-10} (which is the same as calculated K_{eff}=2.00×10^{-10}), for $s = 0.75$ Hyperconv gave K_{F2F}=2.19×10^{-10} (compared to calculated K_{F2F}=2.19×10^{-10}) and for $s = 2$ Hyperconv gave K_{F2F}=3.32×10^{-10} (compared to calculated K_{F2F}=3.33×10^{-10}).

[89] Note that borehole packer testing is likely to reveal heads in the 'active system'.

These are in addition to **boundary nodes**, which are the nodes which form the constant head boundaries.

Modifications to HyperConv were made to allow the number of active and flowing nodes to be counted in particular planar or radial shell intervals. Although any interval length could be chosen, the finest detail possible was sought by fixing the shell intervals to be the same as the nodal spacing. There are a number of ways to present these data, for example by normalising to give a histogram of nodal frequency. There are several possible characteristic numbers. It was found that, on balance, it was useful to leave the data un-normalised in order to examine the unnomalised shapes of the distributions.

13.6 *Ensemble properties*

Because the channel-network is generated stochastically and is here sparsely populated, the shape of the network, the flow structure, the various calculated K and skin values all vary considerably for each random realisation. In order to achieve any clear systematic patterns of behaviour, ensembles of many realisations are needed.

The property being measured (say K, skin, or number of nodes in a given shell) will have a certain distribution over the ensemble. It is cumbersome to present the distribution every time, so the arithmetic mean and SD of the variables are presented as simple summaries. It is worth noting that although distributions have not been exhaustively plotted here, all that were plotted gave relatively smooth monomodal distributions which were either symmetric (and approximately Gaussian) or else mildly skewed. Figure 13.7 gives an example of a histogram for skin, which shows a positive skew.

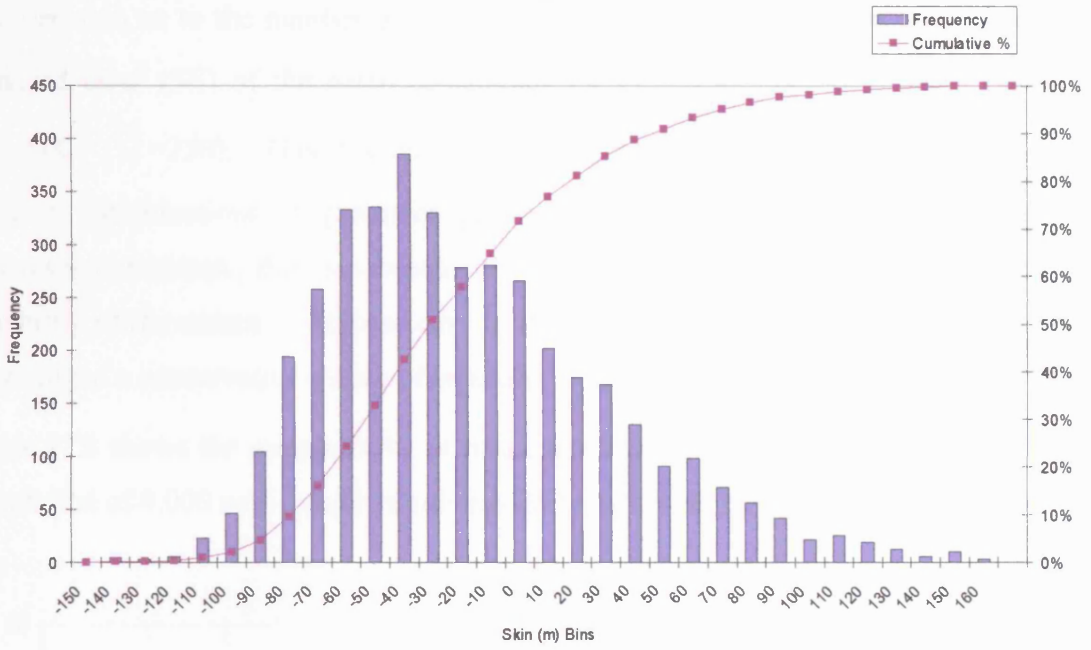


Figure 13.7: Histogram and cumulative frequency of skin based on active nodes s=0, P_A=0.9 and P_N=0.004 (4000 realisations).

Because a real system only exists as a single realisation, the first question that this raises is that of what this 'ensemble' average represents. For K, a spatial average can be shown to tend to the ensemble mean as long as the integral scale, λ, is much smaller than the averaging scale (Gelhar, 1993). Accepting that this is valid is called 'accepting the ergodic hypothesis' in the stochastic hydrogeology literature. Conceptually it is more difficult to see what this means for skin, as it is only an apparent parameter. However, it is essentially derived from the same underlying spatial permeability data, so could be expected to obey similar stochastic rules. The average length scale, $\overline{L} = 10$, is the basic case used in the model runs presented here, which is reasonable given that the F2F system is 59 sub-channels long and the radial system is approximately 30 sub-channels long. The extremes in the distribution of length scales however cannot always be assumed to be much smaller than the domain size. The ergodic assumption in the instance of such channels being generated is questionable. It is better to think of the ensemble results representing the average of datasets at many different locations, but as potentially poor representations of single locations.[90] The mean gap, $\overline{G}$ does not pose any such issues (note that gaps that exceed 59 channels are used without difficulty).

[90] A future study could simulate either much larger systems, or generate much larger networks and select a sub-volume for flow and transport calculations. It is worth noting that the computational expense of doing this would potentially be very high.

The decision as to the number of runs (ensemble size) (n) was based on plots of the standard error (SE) of the mean of samples for the 'base case' of P_A=0.9 ($\overline{L}$=10), P_N=0.004 ($\overline{G}$=250). This ensemble size was then used for multiple model runs at different combinations of generator parameters. Since the SE will vary with the generator parameters, this is subsequently measured for all means as a check on the reliability of the means. The base case is close to a peak in variability of outcome, so encouraged a conservative choice of ensemble size.

Figure 13.8 shows the measured SE of mean skin against 'ideal' SE of mean for a total population of 4,000 runs (which is assumed to be approaching an infinite number).

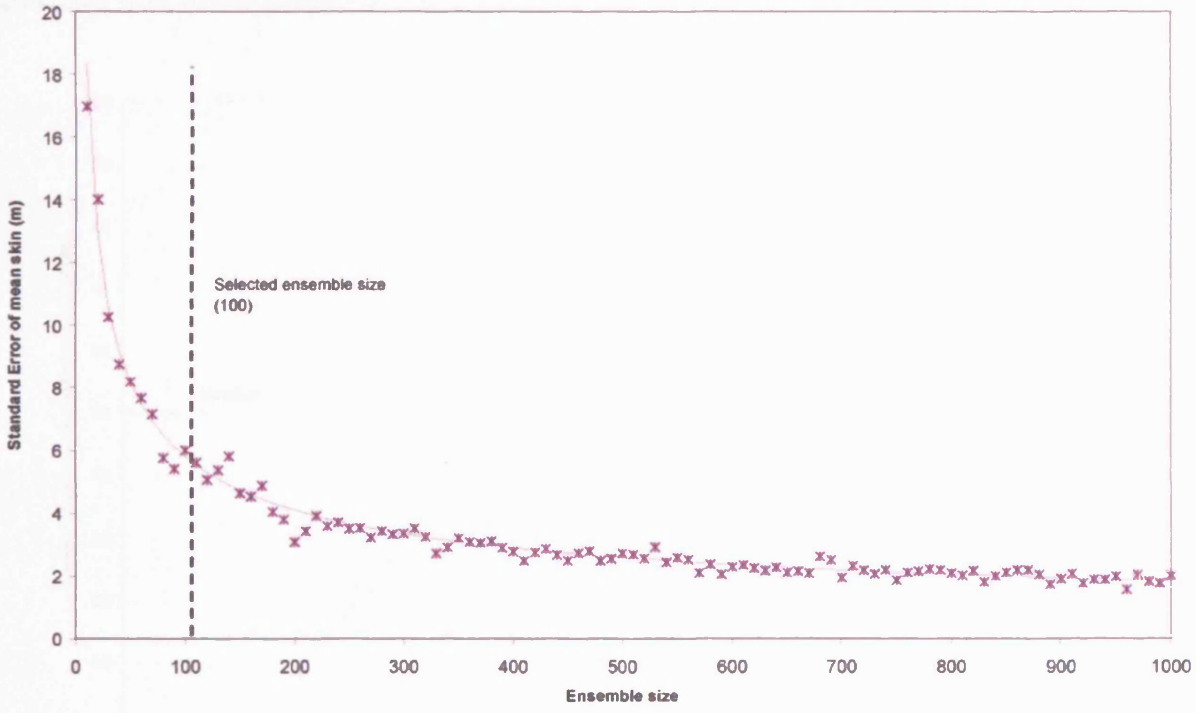


Figure 13.8: SE of mean skin, for P_A=0.9, P_N=0.004, s=0. Line depicts 'ideal' SE $= \frac{\sigma_{4000}}{\sqrt{n}}$.

The 'ideal' SE of the mean, $SE = \frac{\sigma}{\sqrt{n}}$, is computed with the assumption made that σ_{4000} is equal to $\sigma_{real} = \sigma_{\infty}$. It is readily apparent that a substantial reduction in the SE can be achieved by increasing the sample size. A sample size of 100 represents a good compromise between reduced SE and computer run time. This has therefore been adopted as the standard ensemble size, unless otherwise indicated.

The ensemble concept is hereafter used extensively to examine the effect on mean properties by varying network properties and boundary conditions. Percolation is normally discussed in F2F systems and the theory is most developed in relation to infinitely large systems. Here percolation is discussed for the more practical situation of cylindrical flow in a finite system.

Here the percolation threshold, P_C, for the system is defined as the median value of P_{ON} for the distribution of the number of percolating realisations, P_P (i.e. it is the network parameter at which 50% of the ensemble of realisations actually percolated).

Figure 13.9 shows the number percolating for different choices of P_A (i.e. for different average channel lengths, $\overline{L}$) and also for 'classical percolation' (here used to mean where a network is generated with all subchannels having equal probability of being conductions).

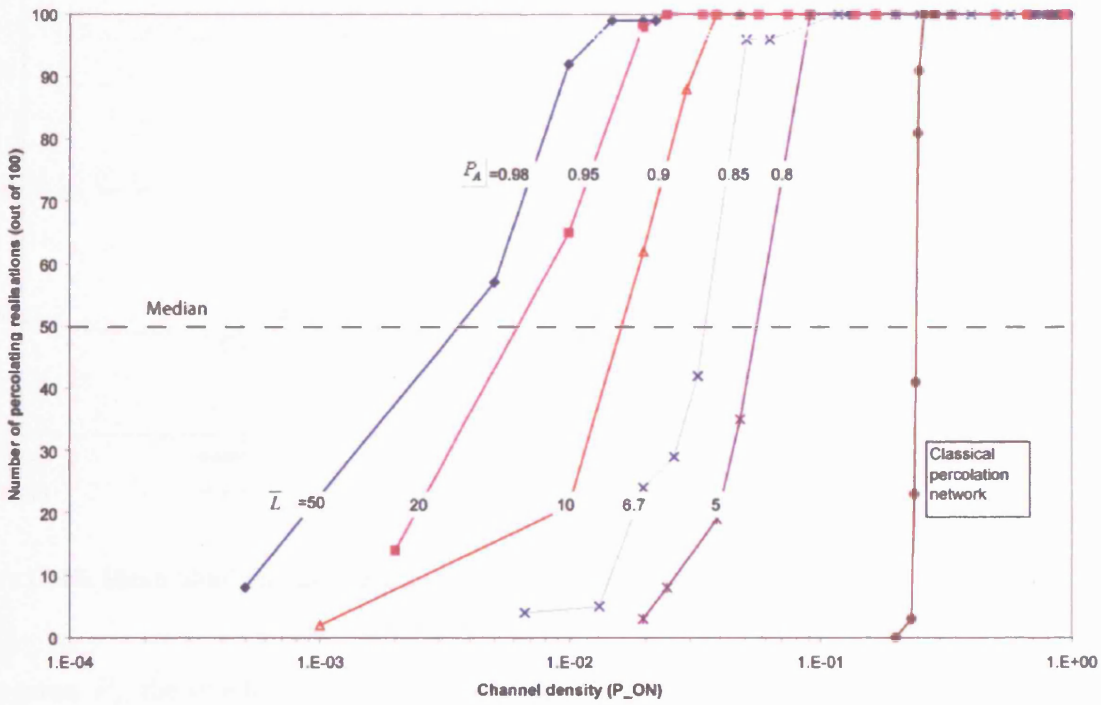


Figure 13.9: Number of percolating realisations, P_P, in cylindrical flow for different P_A and $P_N = 0.9$. P_C for each system is defined by the 50% line. $\overline{L}$ at $P_C = 0.24$ is 1.3.

It is noticeable that the 'classical percolation' is sharply defined and is interpolated to be 0.24 (note that F2F gives 0.25). Note that it is a well-established result for an infinite system that P_C =0.2488 (Sahimi, 1994). The finite boundary and the cylindrical geometry apparently

cause some offset and spreading of the distribution, but only in a minor way. The effect of variable channel lengths is dramatic, both offsetting P_C to occur at lower P_{ON} (i.e. at a less occupied network density) and spreading out the range of P_{ON} at which this occurs.

13.7 *Map of behaviour based on cylindrical skin*

Skin is plotted in Figure 13.10, showing systematic changes in mean skin for different combinations of generator parameters. The underlying distributions of head from which the regression is obtained (along the lines of Figure 13.2) will be explored in Section 13.8.2)

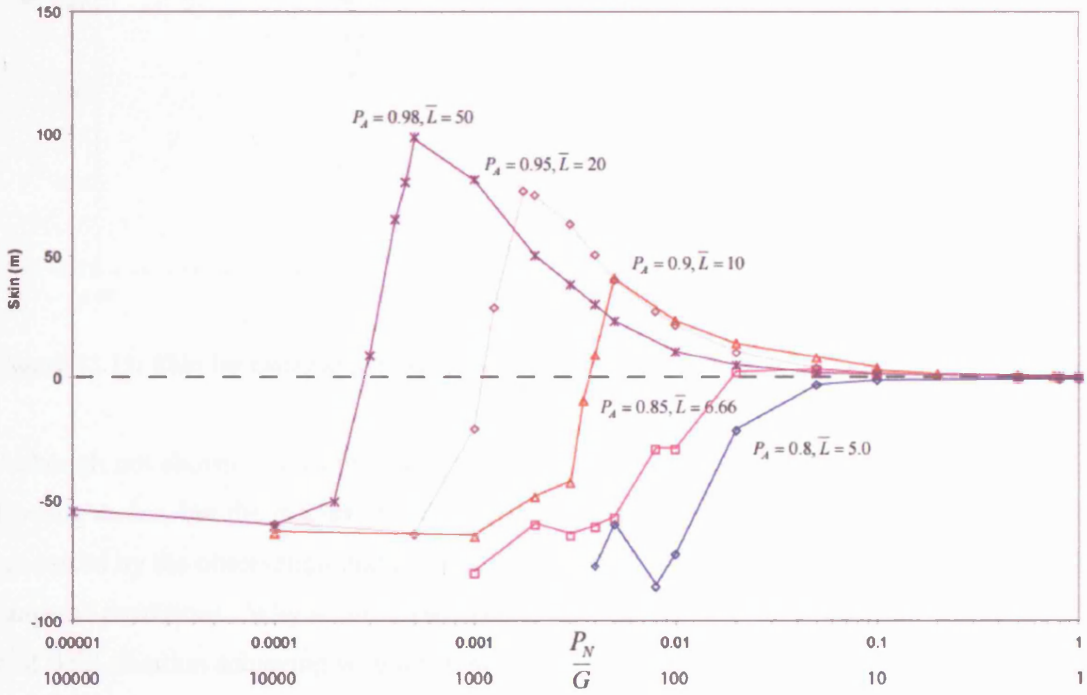


Figure 13.10: Mean 'skin' for varying P_N ($\overline{G}$) and P_A (100 realisations).

For a given P_A, the skin for the realisations that occur below P_C is less than 50 m (remember that P_C is the median P_{ON} of the percolating realisations). As P_N is increased the skin increases steeply up to a peak and then tends to zero as P_N tends to 1. The effect of P_A is to shift the curve to the right of the P_N axis and to reduce the magnitude of the peak. Below P_A=0.8 no positive skin is measured at all. This is true for 'classical' percolation which can be thought of as an extreme end-member in which there is zero correlation (see Figure 13.11). The positive skins observed from field data are therefore either caused by near-drift permeability

reduction, or might be indicative of the sort of increased correlation lengths tested here. They cannot be explained by networks of short channels.

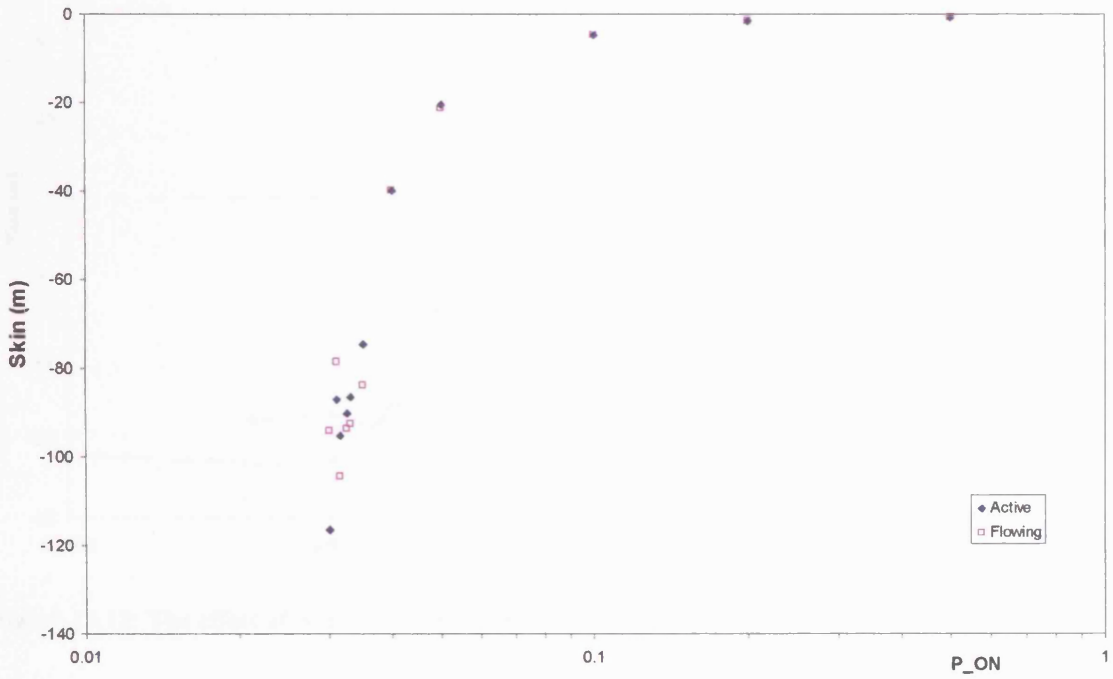


Figure 13.11: Skin for classical percolation network.

Although not shown, the pattern seen in Figure 13.10 for active nodes is also very similar for flowing nodes, but the magnitude of the peaks of positive skin are reduced. Certain questions are raised by the observation that positive skin is therefore only predicted by a relatively narrow range of conditions. Why might these exist in reality? What is the likelihood of a network in a real field situation achieving such a condition?

The SD of $\log_{10}(C_S)$, s, increases the magnitude of skin, apparently slightly reducing the P_N at which transition occurs between very negative and peak positive skin. The effect of s is also to increase the magnitude of this peak of positive skin. The solver experienced convergence problems for non-sparse networks with s>0.75, so the s=2 curve in Figure 13.12 is incomplete.

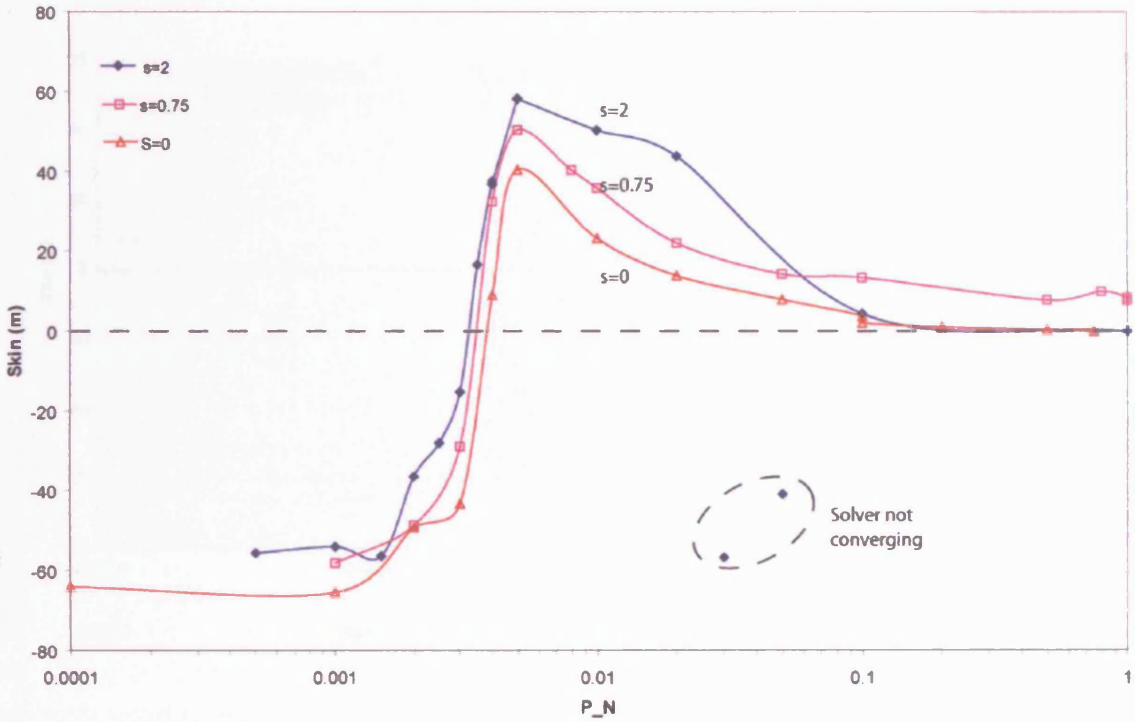


Figure 13.12: The effect of s on skin (100 realisations, P_A=0.9).

Skin therefore varies in a clear family of curves and is controlled by $\overline{L}$ (via P_A) and average gap $\overline{G}$ (via P_N). It is possible to split a particular curve into regions of different skin behaviour for which more detailed examination of the flow and head structure may reveal insights into the causal mechanisms for this pattern. The next sections deal with this more detailed testing.

For the remaining analysis, P_A=0.9 is taken to give a base case, giving an average channel length of 10. As P_N is varied for this base case it is possible to divide the networks generated into 'regions' using the measured skin as an indicator of the underlying behaviour (Figure 13.13). Such a 'map' can be produced for any P_A, but requires a shift in the values of P_N that form the region boundaries. P_N values used to represent each region are therefore taken to be:

Near-critical Percolation: P_N=0.002, $\overline{G}$=500

Transition: P_N=0.004, $\overline{G}$=250

Positive skin regime: P_N=0.008, $\overline{G}$=125

Near-Continuum P_N=0.5, $\overline{G}$=2

clusters holding higher head can exist close to the tunnel, but are connected to the backbone perhaps substantially further away from the tunnel where backbone heads are greater. It is this sort of 'compartmentalisation' that makes near-tunnel observation so unpredictable. This might account for the slight worsening in fit for the near-tunnel points on the 0.008 and 0.004 lines in Figure 13.18.

13.9 *Face to face boundaries*

Having analysed cylindrical boundaries, F2F boundaries are now considered in the same systematic way.

13.9.1 Network distribution (F2F)

Using a large number of realisations has the effect of revealing clear relationships from noise, as illustrated by Figure 13.23 for a 'transition' system (P_N=0.004). Note that the number in each bin (of width one sub-channel) is normalised in order to make the curves overlap and therefore be directly comparable. Examining the 500 run ensemble, it is clear that the number of active nodes increases away from the walls smoothly. A quadratic fits reasonably well. This behaviour appears to reveal a network growing away from a boundary.

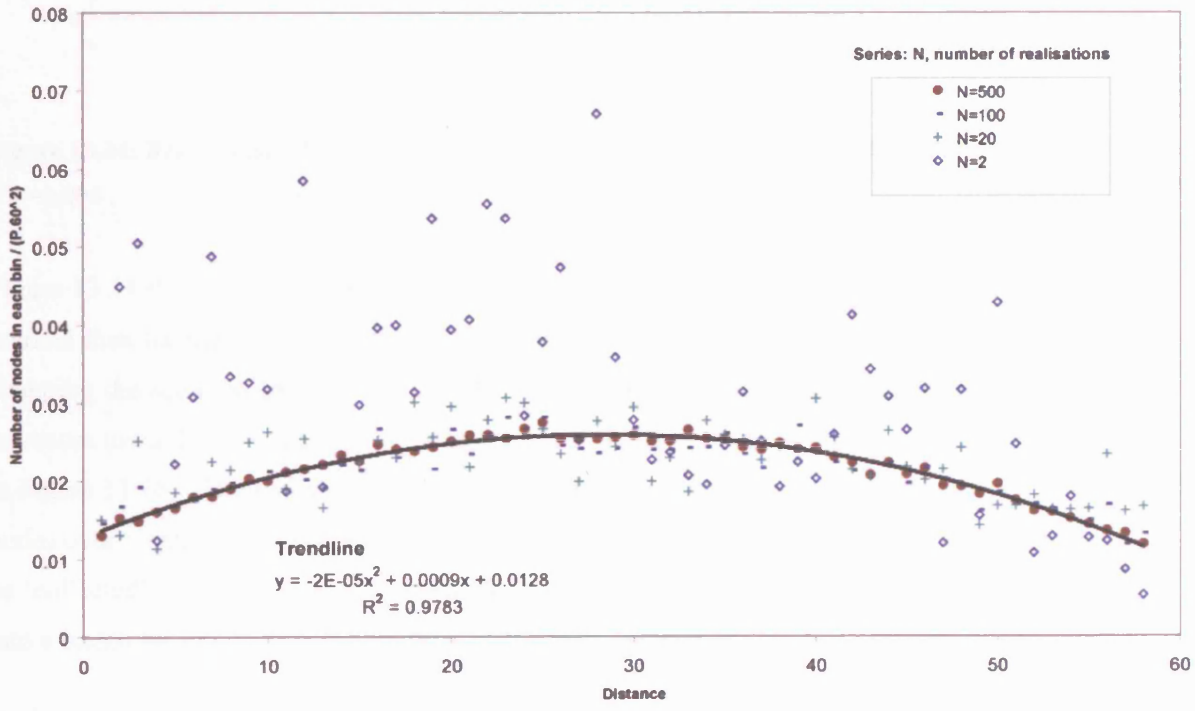


Figure 13.23: F2F – effect of ensemble size on active node distribution for F2F system P_A=0.9, P_N=0.004 and number of nodes is normalised by $P_P.60^2$.

13.11.3 Results

The number and size of pools changes depending on the head drop, Δh, of the bounding waterfalls and the variability of channel conductance. When $s = 0$, for a fully-occupied lattice for $\Delta h > 200/59$, the number of pools would be equal to one. For $\Delta h = 200/59$ the graph would step to 59 pools (each plane of sub-channels aligned perpendicular to the flow). For $\Delta h < 200/59$ the number of pools would remain 59 until the trivial point at $\Delta h = 0$, which would be the total number of sub-channels. The same would occur for a single channel traversing linearly across between the boundaries. If there were a number of traversing channels that did not interconnect, then the step point would be 59 times the number of channels.

For substantially less interconnected systems there is a tendency as Δh is reduced for the number of pools to tend towards the number of active sub-channels. Figure 13.41 compares the number of pools for a given waterfall height for single realisations of the model.

For the 'near-critical percolation' network (P_N=0.002), the number of pools barely changes up to about Δh=3, at which point it suddenly drops. Figure 13.41 shows that s has very little effect on this. For this particular realisation, there were two 'backbones' which connected to both boundaries but not to each other. For a 'near-critical percolation' system this is entirely plausible, because flow is concentrated along single channels with few branching connections. Note that the step point is not far from twice 59, the number that would be expected in the instance of a single channel spanning the boundaries.

For a 'transition' network (P_N=0.004), s appears to be much more important, tending to sharpen the 'elbow' of the curve and reducing the maximum number of pools. At the same time, as s increases the graph is increasingly stepped. This means that the variability in conductance helps to generate some very stable pools, which remain defined for a large range in head. For $s = 1$ with a Δh of 5 m there are 56 pools. However, only six of these pools include more than one node (i.e. 50 pools contain one node). These six pools therefore account for 6,322 out of 6,372 nodes. Packer test head measurement of a system such as this would be most likely to extract data from six discrete pools, each of which is defined by the fact that there is a 5m head difference to the next pool. These pools are highlighted in Figure 13.42. Figure 13.43 shows the head and flow for the same system, demonstrating that the compartments are not necessarily visible by eye.

For a 'positive skin' network (P_N=0.008), the effect of s is again a reduction in the maximum number of pools. This is shown in Figure 13.41. There are no steps in the graph, but instead a continuum of pool sizes with Δh. The lack of steps does not infer a lack of pools. Reducing Δh gradually introduces more pools, in contrast to the distinct thresholds that arise for the P_N=0.002 and 0.004 systems. Packer-test measurement would still potentially experience compartmentalised regions bounded by significant head drops.

The data for these plots could be used to yield average pool sizes. However, this value would be somewhat misleading, because it does not carry information about the distribution of pool sizes.

What is not analysed here is the *range of head* within a pool. It is possible to have a very large pool, within which the head drops very smoothly. The 'pool' conforms to the definition of a pool in as much as there are no waterfalls within it above the threshold head-drop. However, it may not actually exhibit a characteristic signature head or any other property, but may instead be a long sequence of 'rapids'.

The results are clear:

- As the density of the network increases, a larger and more highly interconnected system develops. Rather than control by only a few key chokes, the more dense networks have a much wider range of waterfalls.
- The most striking compartmentalisation occurs for the 'transition' network with larger s.
- By increasing s, even more potential waterfalls are introduced to the system. As the threshold (Δh) of what constitutes a waterfall increases, the size of each pool gradually grows.

It would be worthwhile investigating whether it would be possible to find network parameters that matched observed distributions in flow rates. For example output flow rate could be continuously recorded (logged) in each of the 13 ports over a long period, whilst a fixed-rate pump delivered a constant input flow. Rather than attempting precise matching on a port-by-port basis, examination of what channel network parameters determine the variability in average flow rate between ports, would be very interesting. It would be interesting to see whether a model which included certain special sub-channels which were allowed to 'switch' on and off with a certain periodicity could be configured such that it gave similar flow rate statistics to the proposed flow rate dataset.

If a particle tracking routine was added to HyperConv, then simulations could be done of BTCs at different ports. Careful thought would be needed for how to code the mixing of particles at channel intersections. One system could work by establishing concentration based on counts of non-divisible particles exiting the model (e.g. Johnston et al., 2005). For such particles at each node the code could decide on which sub-channel a particle should next travel down by randomly allocating the particle according to the relative outlet flow rates. It would be possible to attempt calibration of the volumetrically-averaged BTC against a large (e.g. 100 run) ensemble. Alternatively this could be along the lines of the approach used by the MOPOD model (Bloomfield et al., 2005), which is able to track particles, each of a given input mass, through a 2D fracture network model. At each junction between sub-channels (which are called 'fractures' in the MOPOD code) particles are split into smaller sub-particles of mass proportional to the flow rate into each of the adjoining sub-channels. The total travel time of each sub-particle is calculated (the travel time across the sub-channel is added to the total travel time that was achieved upto the previous node). At the outlet of the system the location and total travel time of each sub-particle can be reported.

It is worth considering how dual-porosity and sorption processes might be added. Linear sorption could be applied readily by simply retarding the particle transport velocities. Dual-porosity exchange could be incorporated using a random-walk approach; particles would be allowed walk in the 'matrix' direction (where they do not advect) based on a given residence time distribution function for the block. Alternatively, to incorporate the dual-porosity exchange term, an entirely new transport code could be created which connects together a network of 1D dual-porosity conductors (sub-channels). The transport solution would be obtained by convolution of a series of serially connected transfer functions.

Ultimately, yet more sophisticated models could be constructed that accounted for the production and transport of gas. Capillarity rules could be applied, resulting in a transient

invasion-percolation model, which may have the potential for predicting the sort of flow patterns that arise in the presence of gas. Of particular interest for such a model would be examination of the implications of gas production on solute flushing and on downstream flow-control.

13.13 *Channel network conclusions and recommendations*

Simple orthogonal channel networks with increased channel lengths are demonstrated here to percolate over a wide range of network densities and with a lower median P_C than channels with shorter lengths. Elongate channels may therefore result in sparser networks than 'classical' networks. The longer channels and sparseness give rise to new behaviours.

The Hyperconv model was used to simulate channel networks with channel lengths randomly selected from a geometric distribution. This was done in response to evidence collated by Black et al. (2006) which suggested two key hypotheses about behaviours which could be expected to emerge from a channel network system.

Channel hypothesis 1 anticipated that, as the single channels that connect to tunnel (cylindrical) boundaries connect to other channels, there is a spherical bubble-like growth in the spatial extent of the network away from the tunnel. Measurements of head in a system with spherical D_F would give an apparently 'positive' skin if interpreted by assuming a continuum cylindrical flow system. However, no evidence was found in support of this. Positive skin apparently results from a much more tunnel-local hyperconvergence wherein there is a large head drop across the channels that immediately connect to the tunnel.

The dimensionality of flow away from the tunnel was shown to depend on the density of the network. Less dense (near-critical percolation) gave mainly single-channel flow, thereby giving the opposite effect – negative skin. 'Classical' networks, although denser due to the shorter channels, also gave head profiles that were more characteristic of linear features (i.e. demonstrated hydraulic domination by a single 'backbone').

Channel hypothesis 2 (that there is compartmentalisation) was broadly confirmed by analysis of 'pools' of head.

Given how remarkably revealing the use of histograms of network nodes has been, this spatial frequency approach appears to be very promising and perhaps could be usefully extended to other networks and other network properties (such as flow rates in sub-channels).

Development of an analytical solution that predicts ensemble frequencies of active or flowing nodes with distance remains a challenge. The effect of side boundary conditions requires investigation, perhaps starting with the introduction of periodic side boundaries. The effect of domain scale requires systematic treatment (with much larger networks requiring more powerful computing resources to that currently available on a desktop PC).

There is a need for more modelling work; an analysis of solute transport (perhaps by implementing a particle-tracking code) would be a logical next step. Beyond that, more enhancements could be added, for example, sorption and mechanical dispersion (potentially spatially distributed per channel or sub-channel).

14 Conclusions

This thesis is founded on the challenges and questions posed by the analysis of tracer and flushing data from wastes (**Part II, Chapters 9-12**). The **applied aim** of this research has been to be able to improve the ability to predict contaminant flushing from wastes.

Establishing the dominant processes and the concomitant distributions of parameters based on existing datasets is shown to be a difficult task. However, the dual-porosity mechanism is demonstrated to be important and the block time (t_{cb}) is shown to be of the order of 20 days for a tracer experiment in the Pitsea compression cell and of the order of one day for a tracer experiment in the laboratory. Based on this work, specific improvements for transport experiments in wastes are recommended in **Chapter 12.**

Part I of the thesis examines this problem of diagnosis of process in a more theoretical way, away from the complicating issues that arise for particular datasets. As this analysis is generic, it contributes to a **wider aim** of improving prediction of contaminant flushing from highly heterogeneous materials.

This theoretical analysis questions what might be perceived to be a *de facto* 'conventional' methodology. Relatively limited progress towards the aims of this thesis would have been made by selecting a single model (for example an advection dispersion model with a first order transfer term), fitting it using a gradient method and discussing either the visual fit to the BTC and/or the r^2, then tabulating the parameters and their respective confidence intervals as obtained from the covariance matrix. All such functionality is available within the near-standard CXTFIT code. However, this thesis argues that several steps in such an approach should be questioned and improved. These are set out below:

- Use of a single model or a very narrow selection of models is an inadequate way to test conceptual uncertainty that may have a dominant impact on predictive uncertainty.
- The use of weights in least-square fits needs careful consideration. It is suggested that weighting is related to the structure of the noise model. This in turn requires that there is deeper investigation into the structure of the deviations between the data and models and the various contributions to noise.
- The use of r^2 for confrontation of BTC data with models requires considerable care. In particular what r^2 constitutes a 'good' or a 'poor' fit is not a fixed range of values, but will be specific to the nature of how the model and data differ and what the noise level is.

- Confidence intervals based on the covariance matrix (or Hessian) for transport models that include the dual-porosity process, unless they deviate by a very small amount from the model, should not be relied on.
- The meaning of first order (MIM) parameters and their relation to physically-based models (such as the single block-sized dual-porosity model) requires subtle interpretation. Unless the geometry is well-known it would be a mistake to assume a particular regular geometrical configuration is 'correct'.

Well-known robust and parsimonious process models are examined thoroughly in the thesis. It is important to know the range of each model's behaviour, particularly when parameters are allowed to vary widely. When it comes to examining new media, such as wastes, it is not necessarily obvious *a priori* how to constrain parameters. Furthermore, what constitutes a model's 'behaviour' depends upon what measure is used to examine the model. Three measures are employed simultaneously here (late-time gradients, the method of moments and least-squares fitting), in order to avoid the arbitrariness of adopting a single measure and in order to test to what extent findings based on one measure can be extended to other measures. Simultaneous examination, in the manner illustrated in this thesis, of multiple models by multiple methods is recommended as a more rigorous approach to the examination of transport processes than unsatisfactory serial Popperian elimination of models using arbitrary models and arbitrary rejection criteria followed by subsequent adoption of a new, subjectively-chosen, 'pet hypothesis'.

Loosely-phrased 'rules of thumb' should be applied with caution. It is more satisfactory to test the precise predictions of a specific mathematical model than to make loose 'intuitive' statements and assume that they apply to particular models. Statements such as:

- 'a BTC peak earlier than t_a is indicative of preferential flow';
- 'the early-time data characterises mechanical dispersion';
- 'tests to establish the dual porosity block time must take of the order of the block time';

are not necessarily correct. Also problematic are some attempts to 'logically falsify' models. Examples of erroneous falsification include eliminating a wider hypothesis about the physics of the system based on the failure of a particular mathematical model to match data (despite the fact that the model does not encompass the entire range of possibility contained within the hypothesis).

The study of transport in wastes is at a very interesting and important stage, since the dominant processes are not known (and may vary from site to site and over time). As the science is relatively immature there is not too much dogma about how the system is believed to behave.

Nonetheless, now is a time of rapid import of mass-transport models from other sub-surface literatures. The organisation and appraisal of the range of available models in this thesis (in Chapter 9) is therefore timely. This encourages a more rigorous scientific approach than unquestioningly adopting familiar or convenient models.

The complexity and to some extent inadequacy of waste data, the pressing need for prediction and pressure to identify parameters and processes has stimulated two key strands of work:

1. How to diagnose processes from existing data and how to devise new experiments that are more discriminating.
2. The need to find coherent ways to explore a much wider range of 'model-space' in the light of on-going conceptual uncertainty.

The work on the first strand (diagnosis) potentially forms the back-bone of a future expert system for selecting or weighting process models for use in prediction, based on comparison of models with data. **Chapter 8** summarises the diagnostic traits of lumped and locally stochastic models in terms of late-time concentration gradients, moments and the appearance of breakthrough curves.

The second strand required a coherent way to organise model space and then to populate it with practical, simple models. Predictions based on large ensembles of predictions, weighted by the relative performance of model-parameter combinations are anticipated to be a very important future approach. Such methodologies will need to draw on 'suites' of models, such as those proposed here, covering a wide conceptual space in an organised way.

This diagnostic analytical methodology (and hierarchical approach to BTC analysis) that was developed in Part I has been applied to two waste datasets at laboratory and lysimeter scales (Chapters 10 and 11). This provided a systematic way to structure the analysis and was particularly effective at illuminating where experiments could be enhanced in the future. Given the level of improvement, which appears to be both necessary and possible, Chapter 12 was entirely devoted to elucidating possible future experimental strategies.

These datasets proved to be immensely interesting and valuable beyond their 'path finding' role, revealing key insights at a range of different 'levels'. The majority of the findings were made at levels below the 'modelling level' (level 4) on the analysis hierarchy. This is neither inadequate nor disappointing, since even the connectivity between two points can be a non-trivial matter in highly heterergeneous media. It has been argued that although a limited amount of modelling is

presented here, improvements in experimental methodology are required in order to obtain less ambiguous results.

What is particularly emphasised for the future is an open-minded and 'fertile' approach to conceptual models. Rather than stagnating on 'conventional' approaches it will be important to explore the impact on prediction of new, plausible transport models. An entire chapter is dedicated to one such 'new entrant' to the pool of models – a channel network model. By organising model-space along a 'heterogeneity' axis it remains possible to sample a wide spectrum of approaches without needing to add complexity (i.e. more parameters or assumptions). The channel network model exemplifies this by having a very simple network generator. By generating ensembles of flow realisations, specific characteristics of sparse channel networks (where there are long, thin channels) are examined. A 'map' is drawn of the conditions under which so-called 'positive' skins would be observed by pumping tests from tunnels if the real system behaved like such a network. There is considerable structure to the spatial frequency distribution of channels (and the nodes through which these channels connect). Such structure, although not previously reported in the literature, is shown to extend to 'classical percolation networks'. Simple channel-work models such as these are able to account for 'head compartmentalisation', positive skin and infrequent and unpredictable tunnel inlet point flows.

Recommendations for future work are made throughout the thesis. A specific list of recommendations appears in **Chapter 12**, in terms of how future waste experiments could be run.

In conclusion, the central theme of this thesis is concerned with how to model and analyse transport data in the form of tracer breakthrough curves and flushing concentrations in hitherto poorly characterised and highly heterogeneous systems. The philosophy here is to choose minimally parameterised models, but to test conceptual possibility as widely as possible. An honest appraisal of the state of knowledge of the system is more likely to result if many models are considered as multiple working hypotheses. It is possible to apply diagnostic logic to the simplest lumped models, but for the wider spectrum of possible models it may need to be accepted that no single 'correct' model will ever be identified and prediction may have to be based on a suite of possible models (which can be added to at any time). This encourages future work to be less constrained and more creative conceptually and to move away from sterile and complacent over-use of inadequate models.

References

Abbaspour, C. A., M. T. van Genutchen, R. Schulin, and E. Schlappi. 1997. A Sequential Uncertainty Domain Inverse Procedure for Estimating Subsurface Flow and Transport Parameters. *Water Resources Research,* 33:1879-1892.

Abbaspour, K. C., C. A. Johnson, and M. T. van Genuchten. 2004. Estimating Uncertain Flow and Transport Parameters Using a Sequential Uncertainty Fitting Procedure. *Vadose Zone Journal,* 3:1340-1352.

Abbaspour, K. C., M. A. Sonnleitner, and R. Schulin. 1999. Uncertainty in Estimation of Soil Hydraulic Parameters by Inverse Modelling: Example Lysimeter Experiments. *Soil Sci. Soc. Am. J.,* 63:501-509.

Abelin, H., L. Birgersson, L. Moreno, H. Widen, T. Agren, and I. Neretnieks. 1991. A Large Scale Flow and Tracer Experiment in Granite 2: Results and Interpretation. *Water Resources Research,* 37:3119-3135.

Abramowitz, M., and I. Stegun, A. 1968. *Handbook of Mathematical Functions.* Dover Publications Inc., New York.

Adar, E. M., S. P. Neuman, and D. A. Woolhiser. 1988. Estimation of Spatial Recharge Distribution Using Environmental Isotopes and Hydrochemical Data, I Mathematical Model and Application to Synthetic Data. *Journal of Hydrology,* 97:251-277.

Akaike, H. 1973. Information Theory as an Extension of the Maximum Likelihood Principle. *in* B. N. Petrov, editor. Second International Symposium on Information Theory, Akademiai Kiado, Budapest.

Altmann-Dieses, A. E., J. P. Schlöder, and H. Georg Bock. 2002. Optimal Experimental Design for Parameter Estimation in Column Outflow Experiments. *Water Resources Research,* 38:4-1 - 4-11.

Anderson, M. P., and W. W. Woessner. 1992. *Applied Groundwater Modeling : Simulation of Flow and Advective Transport.* Academic Press, San Diego.

Appelo, C. A. J., and D. Postma. 1994. *Geochemistry, Groundwater and Pollution.* AA Balkema, Rotterdam.

Aris, R. 1975. *The Mathematics of Diffusion and Reaction in Permeable Catalysts, Vols I and II.* Oxford University Press, Oxford.

Aris, R. 1956. On the Dispersion of a Solute in a Fluid Flowing through a Tube. *Proc. Royal Soc. A,* 235:67-77.

Bai, M., D. Elsworth, and J.-C. Roegiers. 1993. Multiporosity/Multipermeability Approach to the Simulation of Naturally Fractured Reservoirs. *Water Resources Research,* 29:1621-1633.

Bajracharya, K., and D. A. Barry. 1997. Nonequilibrium Solute Transport Parameters and Their Physical Significance: Numerical and Experimental Results. *Journal of Contaminant Hydrology,* 24:185-204.

Baranblatt, G. I., I. P. Zheltof, and I. N. Kochina. 1960. Basic Concepts in the Theory of Seepage of Homogeneous Liquids in Fissured Rocks (Strata). *PMM,* 24:852-864.

Bardsley, W. E. 2003. Temporal Moments of a Tracer Pulse in a Perfectly Parallel Flow System. *Advances in Water Resources,* 26:599-607.

Barker, J. A. 1982. Laplace Transform Solutions for Solute Transport in Fissured Aquifers. *Advances in Water Resources,* 5:98-104.

Barker, J. A. 1985a. Block-Geometry Functions Characterizing Transport in Densely Fissured Media. *Journal of Hydrology,* 77:263-279.

Barker, J. A. 1985b. Modelling the Effects of Matrix Diffusion on Transport in Densely Fissured Media. *in* Hydrogeology in the Service of Man, Memoirs of the 18th Congress of the International Association of Hydrogeologists, Cambridge.

Barker, J. A. 1988. A Generalized Radial Flow Model for Hydraulic Tests in Fractured Rock. *Water Resources Research,* 24:1796-1804.

Barker, J. A. 1991. Transport in Fractured Rock. in R. A. Downing and W. B. Wilkinson, editors. *Applied Groundwater Hydrology* Pages 199-216. Clarendon, Oxford.

Barker, J. A., and D. G. Kinniburgh. 1994. *Groundwater Modelling and Modelling Methodology.* Draft R&D Note 295/19/A, National Rivers Authority.

Barker, J. A., D. G. Kinniburgh, D. M. J. Macdonald, D. J. Allen, and A. M. MacDonald. 1995. *Groundwater Modelling and Modelling Methodology.* R&D Project Record 295/21/A, National Rivers Authority, Bristol.

Barker, J. A., T. Wright, and B. Fretwell. 2000. A Pulsed-Velocity Method of Double Porosity Solute Transport Modelling. *in* Dassargues, editor. International Conference on Tracers and Modelling in Hydrogeology. IAHS, Liège, Belgium.

Baveye, P., and G. Sposito. 1984. The Operational Significance of the Continuum Hypothesis in the Theory of Water Movement through Soils and Aquifers. *Water Resources Research,* 20:521-530.

Baveye, P. 2004. The Emergence of a New Kind of Relativism in Environmental Modelling: A Commentary. *Proc. Royal. Soc. Lond. A,* 460:2141-2146.

Bear, J. 1960. *The Transition Zone between Fresh and Salt Waters in Coastal Aquifers.* University of California, Berkeley.

Bear, J. 1972. *Dynamics of Fluids in Porous Media.* American Elsevier Publishing Company Inc, Toronto.

Beaven, R., and A. Hudson. 2002. Description of a Tracer Experiment for Evaluating Flows through Waste. *in* Waste 2002: Integrated Waste Management and Pollution Control: Research, Policy and Practice., Stratford-Upon-Avon, Warwickshire.

Beaven, R. P. 2000. *The Hydrogeological and Geotechnical Properties of Household Waste in Relation to Sustainable Landfilling.* PhD. University of London, London.

Beaven, R. P., J. A. Barker, and A. Hudson. 2003. Description of a Tracer Test through Waste and Application of a Double Porosity Model. *in* Sardinia 2003: Proceedings of the Ninth International Waste Management and Landfill Symposium, Margherita di Pula, Cagliari, Sardinia, Italy.

Beaven, R. P., J. A. Barker, and N. D. Woodman. 2005a. End-Member Flushing Models for 'Saturated' Waste. *in* S. Cossu, editor. Sardinia 2005 Tenth International Waste Management and Landfill Symposium, S Margherita di Pula, Cagliari, Sardinia, Italy.

Beaven, R. P., L. Dollar, L. Oni, and N. D. Woodman. 2005b. A Laboratory Scale Saturated and Unsaturated Tracer Test through Waste. *in* International Workshop, Hydro-Physico Mechanics of Landfills, Lirigm,University of Grenoble.

Becker, M. W., and R. J. Charbeneau. 2000. First-Passage-Time Transfer Functions for Groundwater Tracer Tests Conducted in Radially Convergent Flow. *Journal of Contaminant Hydrology,* 40:299-310.

Becker, M. W., and T. B. Coplen. 2001. Technical Note: Use of Deuterated Water as a Conservative Artificial Groundwater Tracer. *Hydrogeology Journal,* 9:512-516.

Becker, M. W., and A. M. Shapiro. 2000. Tracer Transport in Fractured Crystalline Rock: Evidence of Nondiffusive Breakthrough Tailing. *Water Resources Research,* 36:1677-1686.

Bendz, D., and V. P. Singh. 1999. Solute Transport under Steady and Transient Conditions in Biodegraded Municipal Solid Waste. *Water Resources Research,* 35:2333-2345.

Bendz, D., V. P. Singh, and M. Akesson. 1997. Accumulation of Water and Generation of Leachate in a Young Landfill. *Journal of Hydrology,* 203:1-10.

Bendz, D., V. P. Singh, H. Rosqvist, and L. Bengtsson. 1998. Kinematic Wave Model for Water Movement in Municipal Solid Waste. *Water Resources Research,* 34:2963-2970.

Bengtsson, L., D. Bendz, W. Hogland, H. Rosqvist, and M. Akesson. 1994. Water Balance for Landfills of Different Age. *Journal of Hydrology,* 158:203-217.

Benjamin, J. R., and C. A. Cornell. 1970. *Probability, Statistics and Decision for Civil Engineers.* McGraw-Hill, New York.

Berglund, S., and V. Cvetkovic. 1996. Contaminant Displacement in Aquifers: Coupled Effects of Flow Heterogeneity and Nonlinear Sorbtion. *Water Resources Research,* 32:23-32.

Berkowitz, B. 2002. Characterising Flow and Transport in Fractured Geological Media: A Review. *Advances in Water Resources,* 25:861-884.

Berkowitz, B., and I. Balberg. 1993. Percolation Theory and Its Application to Groundwater Hydrology. *Water Resources Research,* 29:775-794.

Berkowitz, B., J. Bear, and C. Braester. 1988. Continuum Models for Contaminant Transport in Fractured Porous Formations. *Water Resources Research,* 24:1225-1236.

Berkowitz, B., and H. Scher. 1998. Theory of Anomalous Transport in Random Fracture Networks. *Physical Review E,* 57:5858-5869.

Berkowitz, B., and C. Braester. 1991. Dispersion in Sub-Representative Elementary Volume Fracture Networks: Percolation Theory and Random Walk Approaches. *Water Resources Research,* 27:3159-3164.

Berkowitz, B., and H. Scher. 1995. On Characterization of Anomalous Dispersion in Porous and Fractured Media. *Water Resources Research,* 31:1461-1466.

Berkowitz, B., and G. Margolin. 2004. Continuous Time Random Walks Revisited: First Passage Time and Spatial Distributions. *Physica A,* 334:46-66.

Berkowitz, B., and H. Scher. 2001. The Role of Probabilistic Approaches to Transport Theory in Heterogeneous Media. *Transport in Porous Media,* 42.

Beven, K. 1996. Equifinality and Uncertainty in Geomorphological Modelling. *in* B. L. Rhoads and C. E. Thorn, editors. The Scientific Nature of Geomorphology: Proceedings of the 27th Birmingham symposium in Geomorphology, Birmingham.

Beven, K. 2002. Towards a Coherent Philosophy for Modelling the Environment. *Proc. R. Soc. Lond. A,* 458:2465-2484.

Beven, K. 2005. A Manifesto for the Equifinality Thesis. *Journal of Hydrology,* 20:18-36.

Beven, K., and A. Binley. 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction. *Hydrological processes,* 6:279-298.

Beven, K., and J. Freer. 2001. Equifinality, Data Assimilation, and Uncertainty Estimation in Mechanistic Modelling of Complex Environmental Systems Using the GLUE Methodology. *Journal of Hydrology,* 249:11-29.

Beven, K., and P. Germann. 1981. Water Flow in Soil Macropores II. A Combined Flow Model. *Journal of Soil Science,* 32:15-29.

Bibby, R. 1981. Mass Transport of Solutes in Dual-Porosity Media. *Water Resources Research,* 17:1075-1081.

Birgersson, I., I. Neretnieks, H. Widen, and T. Agren. 1994. In-Situ Tracer Experiments in Fracture Zones and Averagely Fractured Rock. Pages 171-194 *in* Proc. of the 4th International NEA/SKB Symposium, In Situ Experiments at the Stripa Mine. OECD, Stockholm.

Black, J., P. Robinson, and J. A. Barker. 2005. *A Preliminary Investigation of the Concept of 'Hyper-Convergence' Using 'Sparse' Channel Networks. Report to SKB.* ISS 2003-3 B2, In Situ Solutions.

Black, J. H., J. A. Barker, and N. D. Woodman. 2006. *An Investigation of 'Sparse Channel Networks': Characteristic Behaviours and Their Causes.* ISS 2006-I, In Situ Solutions.

Blakey, N. 1982. Infiltration and Adsorption of Water by Domesti Wastes in Landfills, Research Carried out by Wrc, Paper Presented at Harwll Landfill Leachate Symposium, Uk. *in,* UK Atomic Energy Authority, Oxon, England.

Blakey, N. C., K. Blackmore, and L. Clarke. 1998. *Application of Tracer Studies for Monitoring Leachate Recirculation in Landfills.* CWM 171/98, Environment Agency, Report CWM 171/98.

Blight, G. E., J. M. Ball, and J. J. Blight. 1992. Moisture and Suction in Sanitary Landfills in Semiarid Areas. *Journal of Environmental Engineering,* 118:865-877.

Bloomfield, J.P., Barker, J.A. and Robinson, N., 2005. Modelling fracture porosity development using simple growth laws. Ground Water, 44(3): 314-325.

Bodin, J., F. Delay, and G. De Marsily. 2003a. Solute Transport in a Single Fracture with Negligible Matrix Permeability: 1. Fundamental Mechanisms. *Hydrogeology Journal,* 11:418-433.

Bodin, J., F. Delay, and G. De Marsily. 2003b. Solute Transport in a Single Fracture with Negligible Matrix Permeability: 2. Mathematical Formalism. *Hydrogeology Journal,* 11:434-454

Bonnet, E., O. Bour, N. E. Odling, P. Davy, I. Main, P. Cowie, and B. Berkovitz. 2001. Scaling of Fracture Systems in Geological Media. *Reviews of Geophysics,* 39:347-383.

Bosma, W. J. P., S. E. A. T. M. van der Zee, and C. J. van Duijn. 1996. Plume Development of a Nonlinearly Adsorbing Solute in Heterogeneous Porous Formations. *Water Resources Research,* 32:1569-1584.

Boyer, Morrison, Pavone, and Schwerer. 1982. *Methane Modelling - Predicting the Inflow of Methane Gas into Coal.* DOE final report De-Ac22-80pc30123.

Brenner, H. 1962. The Diffusion Model of Longitudinal Mixing in Beds of Finite Length. Numerical Values. *Chem. Eng. Sci.,* 17.

Bruggeman, G. A. 1999. *Analytical Solutions of Geohydrological Problems.* Elsevier, Oxford.

Brusseau, M. L. 1995. The Effect of Nonlinear Sorption on Transformation of Contaminants During Transport in Porous Media. *Journal of Contaminant Hydrology,* 17:277-291.

Brusseau, M. L., Q. Hu, and R. Srivastava. 1997. Using Flow Interruption to Identify Factors Causing Nonideal Contaminant Transport. *Journal of Contaminant Hydrology,* 24.

Cacas, M. C., E. Ledoux, G. DeMarsily, and B. Tillie. 1990. Modeling Fracture Flow with a Stochastic Discrete Fracture Network: Calibration and Validation. 1. The Flow Model. *Water Resources Research,* 26:479-489.

Carneiro, J. 2000. *A Study on New Approaches for Delineating Groundwater Protection Zones in Fractured-Rock Aquifers.* PhD. UCL, London.

Carrera, J. 1993. An Overview of Uncertainties in Modelling Groundwater Solute Transport. *Journal of Contaminant Hydrology.* 13:23-48.

Carrera, J., A. Alcokea, A. Medina, J. Hidalgo, and L. Slooten, J. 2005. Inverse Problem in Hydrogeology. *Hydrogeology Journal,* 13:206-222.

Carrera, J., and L. Martinez-Landa. 2000. Mixed Discrete-Continuum Models: A Summary of Experiences in Test Interpretation and Model Prediction. in B. Faybishenko, P. A. Witherspoon, and S. M. Benson, editors. *Dynamics of Fluids in Fractured Rock.* American Geophysical Union, Washinton DC.

Carrera, J., X. Sanchez-Vila, I. Benet, A. Medina, G. Galarza, and J. Guimera. 1998. On Matrix Diffusion: Formulations, Solution Methods and Qualitative Effects. *Hydrogeology Journal,* 6:178-190.

Carslaw, H. S., and J. C. Jaeger. 1959. *Conduction of Heat in Solids.* Oxford Science Publications, Oxford.

Chamberlain, T. C. 1897. The Method of Multiple Working Hypotheses. *Journal of Geology,* 5.

Cirpka, O. A., and P. K. Kitanidis. 2000a. An Advective-Dispersive Stream Tube Approach for the Transfer of Conservative-Tracer Data to Reactive Transport. *Water Resources Research,* 36:1209-1220.

Cirpka, O. A., and P. K. Kitanidis. 2000b. Characterisation of Mixing and Dilution in Heterogeneous Aquifers by Means of Local Temporal Moments. *Water Resources Research,* 36:1221-1236.

Cleary, R. W., and D. D. Adrian. 1973. Analytical Solution of the Convective-Dispersive Equation for Cation Adsorption in Soils. *Soil Sci. Soc. Am. Proc.,* 37:197-199.

Cornaton, F., and P. Perrochet. 2002. Analytical 1D Dual-Porosity Equivalent Solutions to 3D Discrete Single-Continuum Models. Application to Karstic Spring Hydrogrpah Modelling. *Journal of Hydrology,* 262:165-176.

Coy, V. 2001. *A Study of Matrix Block Geometry with Respect to Dual-Porosity Models and Tracer Test Analysis.* MSc. UCL.

Crank, J. 1975. *The Mathematics of Diffusion.* Clarendon Press, Oxford.

Cunningham, J. A., and P. V. Roberts. 1998. Use of Temporal Moments to Investigate the Effects of Nonuniform Grain-Size Distribution on the Transport of Sorbing Solutes. *Water Resources Research,* 34:1415-1425.

Cunningham, J. A., C. J. Werth, M. Reinhard, and P. V. Roberts. 1997. Effects of Grain-Scale Mass Transfer on the Transport of Volatile Organics through Sediments. 1. Model Development. *Water Resources Research,* 33:2713-2726.

Cushman, J. H. 1990. *Dynamics of Fluids in Hierarchical Porous Media.* Academic Press, London.

Cvetkovic, V., and G. Dagan. 1994. Transport of Kinetically Sorbing Solute by Steady Random Velocity in Heterogeneous Porous Formations. *J. Fluid Mech.,* 265:189-215.

Cvetkovic, V., G. Dagan, and H. Cheng. 1998. Contaminant Transport in Aquifers with Spatially Variable Hydraulic and Sorption Properties. *Proc. R. Soc. Lond. A,* 454:2173-2207.

Cvetkovic, V., A. Shapiro, and G. Dagan. 1992. A Solute Flux Approach to Transport in Heterogeneous Formations 2. Uncertainty Analysis. *Water Resources Research,* 28:1377-1388.

Dagan, G. 1988. Time-Dependent Macrodispersion for Solute Transport in Anisotropic Heterogenous Aquifer. *Water Resources Research,* 24:1491-1500.

Dagan, G., and V. Cvetkovic. 1996. Reactive Transport and Immiscible Flow in Geological Media. I. General Therory. *Proc. R. Soc. Lond. A,* 452:285-301.

Dagan, G., V. Cvetkovic, and A. Shapiro. 1992. A Solute Flux Approach to Transport in Heterogeneous Formations 1. The General Framework. *Water Resources Research,* 28:1369-1376.

Day-Lewis, F. D. 2000. Identifying Fracture-Zone Geometry Using Simulated Annealing and Hydraulic-Connection Data. *Water Resources Research,* 36:1707-1721.

de Dreuzy, J.-R., P. Davy, and O. Bour. 2000. Percolation Parameter and Percolation-Threshold Estimates for Three-Dimensional Random Ellipses with Widely Scattered Distributions of Eccentricity and Size. *Physical Review E,* 62:5948-5952.

Deitsch, J. J., J. A. Smith, M. B. Arnold, and J. Bolus. 1998. Sorption and Desorption Rates of Carbon Tetrachloride and 1,2-Dichlorobenzene to Three Organobentonites and a Natural Peat Soil. *Environmental Science and Technology,* 32:3169-3177.

Delay, F., G. Porel, and O. Banton. 1998. An Approach to Transport in Heterogeneous Porous Media Using the Truncated Temporal Moment Equations: Theory and Numerical Validation. *Transport in Porous Media,* 32:199-232.

Demetracopoulos, A. C., L. Sehayek, and H. Erdogan. 1986. Modeling Leachate Production from Municipal Landfills. *J. of the Environmental Engineering Division,* 112:849-866.

Dershowitz, W., G. Lee, J. Geier, S. Hitchcock, and L. P. P. 1995. FRACMAN, Interactive Discrete Feature Data Analysis, Geometric Modeling, and Exploration Simulation, User Documentation. *in.* Golder Associates Inc, Seattle, Washington.

De Swaan. 1990. Influence of Shape and Skin of Matrix-Rock Blocks on Pressure Transients in Fractured Reservoirs. *SPE Formation Evaluation,* 5:344-352.

Detwiler, R. L., H. Rajaram, and R. J. Glass. 2000. Solute Transport in Variable-Aperture Fractures: An Investigation of the Relative Importance of Taylor Dispersion and Macrodispersion. *Water Resources Research,* 36:1611-1625.

Dimitrova, J., S. Attinger, and W. Kinzelbach. 2002. *in* Metz, editor. TRePro 2002 Modelling of Coupled transport reaction processes, Forschungszentrum Karlsrhr.

Döberl, G., R. Huber, J. Fellner, and P. H. Brunner. 2003. The Heterogeneity of Waste as a Main Factor Influencing the Future Emission Potential of MSW Landfills - a Case Study on the Breitenau-Landfill in Austria. *in* Christensen, editor. Sardinia 2003, Ninth International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Sardinia, Italy.

Doherty, J. 1994. *Pest Model-Independent Parameter Estimation.* Watermark Numerical Computing.

Dykhuizen, R. C. 1990. A New Coupling Term for Dual-Porosity Models. *Water Resources Research,* 26:351-356.

Efron, B., and R. J. Tibshirani. 1993. *An Introduction to the Bootstrap.* Chapman & Hall, London.

Ehrig, H. J. 1983. Quantity and Quality of Sanitary Landfill Leachate. *Waste Management and Research,* 1:53-68.

Einstein, A. 1956. *Investigations on the Theory of the Brownian Movements.* Dover Publications, New York.

El-Fadel, M., A. N. Findikakis, and J. O. Leckie. 1996. Numerical Modelling of Generation and Transport of Gas and Heat in Landfills I. Model Formulation. *Waste Management and Research,* 14:483-504.

El-Fadel, M., A. N. Findikakis, and J. O. Leckie. 1997. Modeling Leachate Generation and Transport in Solid Waste Landfills. *Environmental Technology,* 18:669-686.

Farlow, S. J. 1982. *Partial Differential Equations for Scientists and Engineers.* John Wiley and Sons, Chichester.

Fellner, J., R. Huber, G. Döberl, and P. H. Brunner. 2003. Hydraulics of MSW Landfills and Its Implications for Water Flow Modelling. *in* Christensen, editor. Sardinia 2003, Ninth International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Sardinia, Italy.

Fetter, C. W. 1999. *Contaminant Hydrogeology.* Prentice Hall, New Jersey.

Field, M. 1999. *The QTRACER2 Program for Tracer-Breakthrough Curve Analysis for Tracer Tests in Karstic Aquifers and Other Hydrologic Systems.* EPA/600/R-02/001/May 2002, U.S. E.P.A., Washington.

Flühler, H., W. Durner, and M. Flury. 1996. Lateral Solute Mixing Processes - a Key for Understanding Field-Scale Transport of Water and Solutes. *Geoderma,* 70.

Flury, M., and H. Flühler. 1995. Tracer Characteristics of Brilliant Blue FCF. *Soil Sci. Soc. Am. J.,* 59:22-27.

Fortin. 1997. Rate-Limited Sorption of Simazine in Saturated Soil Columns. *Journal of Contaminant Hydrol.,* 25:219-234.

Freer, J., K. Beven, and B. Ambrose. 1996. Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach. *Water Resources Research,* 32:2161-2173.

Freeze, R. A. 1975. A Stochastic-Conceptual Analysis of One-Dimentional Groundwater Flow in Nonuniform Homogeneous Media. *Water Resources Research,* 11:725-741.

Freeze, R. A., and J. A. Cherry. 1979. *Groundwater.* Prentice Hall, New Jersey.

Fretwell, B. 1999. *Distribution of Contaminants in the Seasonally Unsaturated Zone of the Chalk Aquifer.* PhD. UCL, London.

Gale, J. E., P. A. Witherspoon, C. R. Wilson, and A. Rouleau. 1982. Hydrogeological Characterisation of the Stripa Site. *in* In-situ Experiments in Granite, OECD/NEA Workshop, Stockholm, Sweden.

Garboczi, E. J., K. A. Snyder, J. F. Douglas, and M. F. Thorpe. 1995. Geometrical Percolation Threshold of Overlapping Ellipsoids. *Physical Review E,* 52:819 LP - 828.

Garnier, J. M., N. Crampon, C. Préaux, G. Porel, and M. Vreulx. 1985. Traçage Par ^{13}C, ^{2}H, ^{-}I et Uranine dans la Nappe de la Craie Sénsonienne en Écoulement Radial Convergent (Bètune, France). *Journal of Hydrology,* 78:379-392.

Gaus, I., A. T. Williams, and P. Shand. 2000. *ASR – UK Physical and Geochemical Modelling (Swift-Phreeqc) of British Aquifers for Aquifer Storage and Recovery Purposes. Part 1: Physical Modelling and Geochemical Model Calibration.* BGS Technical Report WD/00/08, British Geological Survey.

Gawande, N. G., D. R. Reinhart, and C. A. L. G. 2005. Landfill MSW Hydraulic Conductivity Estimation Using in Situ Moiture Sensors. *in* R. Cossu and R. Stegmann, editors. Sardinia 2005, Tenth International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Italy.

Gelhar, L. W. 1986. Stochastic Subsurface Hydrology from Theory to Applications. *Water Resources Research,* 22:135S-145S.

Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology.* Prentice-Hall, New Jersey.

Gelhar, L. W., and C. L. Axness. 1983. Three-Dimensional Stochastic Analysis of Macrodispersion in Aquifers. *Water Resources Research,* 19:161-174.

Gerke, H. H., and M. T. van Genutchen. 1993. Evaluation of a First-Order Water Transfer Term for Variably Saturated Dual-Porosity Flow Models. *Water Resources Research,* 29:1225-1238.

Gerke, H. H., and M. T. van Genutchten. 1993. A Dual-Porosity Model for Simulating the Preferential Movement of Water and Solutes in Structured Porous Media. *Water Resources Research,* 29:305-319.

Gershon, N. D., and A. Nir. 1969. Effects of Boundary Conditions on Models of Tracer Distribution in Flow through Porous Mediums. *Water Resources Research,* 5:830-839.

Glass, R. J., M. J. Nicholl, S. E. Pringle, and T. R. Wood. 2002. Unsaturated Flow through a Fracture-Matrix Network: Dynamic Preferential Pathways in Mesoscale Laboratory Experiments. *Water Resources Research,* 38:1281, doi:1210.1029/2001WR001002.

Goltz, M. N., and P. V. Roberts. 1987. Using the Method of Moments to Analyse Three-Dimensional Diffusion-Limited Solute Transport from Temporal and Spatial Perspectives. *Water Resources Research,* 23:1575-1585.

Gomez-Hernandez, J. J., and R. M. Srivastava. 1990. ISIM3D. An Ansi-C Three Dimensional Multiple Indicator Conditional Simulation Program,. *Computers and Geosciences,* 16:395-440.

Gönüllü, M. T. 1994a. Analytical Modelling of Inorganic Contaminants in Leachate. *Waste Management and Research,* 12:339-350.

Gönüllü, M. T. 1994b. Analytical Modelling of Organic Contaminants in Leachate. *Waste Management and Research,* 12:141-150.

Goodman, R. E., D. G. Moye, A. van Schalkwyk, and I. Javandel. 1965. Ground Water Inflows During Tunnel Driving. *Engineering Geology,* 2:39-56.

Greenkorn, R. A., and D. P. Kessler. 1969. Dispersion in Heterogeneous Nonuniform Anisotropic Porous Media. *Ind. Eng. Chem.,* 61:14-32.

Griffioen, J. W. 1998. Suitability of the First-Order Mass Transfer Concept for Describing Cyclic Diffusive Mass Transfer in Stagnant Zones. *Journal of Contaminant Hydrology,* 34:155-165.

Griffioen, J. W., D. A. Barry, and J.-Y. Parlange. 1998. Interpretation of Two-Region Model Parameters. *Water Resources Research,* 34:373-384.

Grove, D. B., and K. G. Stollenwerk. 1984. *Computer Model of One-Dimensional Equilibrium Controlled Sorption Processes.* U.S.G.S. Water Resources Investigation 84-4059, U.S. Geological Survey.

Gupta, H. V., S. Sorooshian, and Yapo. 1998. Toward Improved Calibration of Hydrologic Models: Multiple and Noncommensurable Measures of Information. *Water Resources Research,* 34:751-763.

Guswa, A. J., and D. L. Freyberg. 2002. On Using the Equivalent Conductivity to Characterise Solute Spreading in Environments with Low-Permeability Barriers. *Water Resources Research,* 38.

Gylling, B., L. Birgersson, L. Moreno, and I. Neretnieks. 1998. Analysis of a Long-Term Pumping and Tracer Test Using the Channel Network Model. *Journal of Contaminant Hydrology,* 32:203-222.

Gylling, B., L. Moreno, and I. Neretnieks. 1999. The Channel Network Model - a Tool for Transport Simulation in Fractured Media,. *Groundwater,* 37:367-375.

Gwo, J. P., P. M. Jardine, G. V. Wilson, and G. T. Yeh. 1995. A Multiple Pore-Region Concept to Modeling Mass Transfer in Subsurface Media. *Journal of Hydrology,* 164:217-237.

Hadermann, J., and W. Heer. 1996. The Grimsel (Switzerland) Migration Experiment: Integrating Field Experiments, Laboratory Investigations and Modelling. *Journal of Contaminant Hydrology,* 21:87-100.

Haggerty, R. 1999. *Application of the Multirate Diffusion Approach in Tracer Test Studies at Äspö HRL.* R-99-62, Svensk Kärnbränslehantering AB (SKB).

Haggerty, R., and S. M. Gorelick. 1995. Multiple-Rate Mass Transfer for Modeling Diffusion and Surface Reactions in Media with Pore-Scale Heterogeneity. *Water Resources Research,* 31:2383-2400.

Haggerty, R., and S. M. Gorelick. 1998. Modeling Mass Transfer Processes in Soil Columns with Pore-Scale Heterogeneity. *Soil Sci. Am. J.,* 62:62-74.

Haggerty, R., S. A. McKenna, and L. C. Meigs. 2000. On the Late-Time Behavior of Tracer Test Breakthrough Curves. *Water Resources Research,* 36:3467-3479.

Han, B., B. Jafarpour, V. Gallagher, P. T. Imhoff, P. C. Chiu, and D. A. Fluman. 2006. Measuring Seasonal Variations of Moisture in a Landfill with the Partitioning Gas Tracer Test. *Waste Management,* 26:344-355.

Hannan, E. J., and B. G. Quinn. 1979. The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society,* Series B41:190-195.

Harris, M. R. R. 1979. *A Study on the Behaviour of Refuse as a Landfill Material.* PhD. Portsmouth Polytechnic.

Hartley, L. J. 1998. *NAPSAC (Release 4.1) Technical Summary Document.* AEA-D&R-0271, AEA Technology, Harwell, Oxon.

Harvey, C., and S. M. Gorelick. 2000. Rate-Limited Mass Transfer or Macrodispersion: Which Dominates Plume Evolution at the Macrodispersion Experiment (MADE) Site? *Water Resources Research,* 36:637-650.

Harvey, C. F., R. Haggerty, and S. M. Gorelick. 1994. Aquifer Remediation: A Method for Estimating Mass Transfer Rate Coefficients and an Evaluation of Pulsed Pumping. *Water Resources Research,* 30:1979-1991.

Hayward, D. O., and B. M. W. Trapnell. 1964. *Chemisorption.* Butterworths, London.

Herrera, I. 1974. Integrodifferential Equations for Systems of Leaky Aquifers and Applications. 2. Error Analysis of Approximate Theories. *Water Resources Research,* 10:811-821.

Herrera, I. 1976. A Review of the Integrodifferential Equations Approach to Leaky Aquifer Mechanics. in Saleem, editor. *Advances in Groundwater Hydrology.* American Water Resources Association, Minneapolis.

Herrera, I., and L. Rodarte. 1973. Integrodifferential Equations for Systems of Leaky Aquifers and Applications. 1. The Nature of Approximate Theories. *Water Resources Research,* 9:995-1005.

Herrera, I., and R. Yates. 1977. Integrodifferential Equations for Systems of Leaky Aquifers and Applications. 3. A Numerical Method of Unlimited Applicability. *Water Resources Research,* 13:725-732.

Hilborn, R., and M. Mangel. 1997. *The Ecological Detective. Confronting Models with Data.* Princeton University Press, Chichester.

Hinz, C., L. A. Gaston, and H. M. Selim. 1994. Effect of Sorption Isotherm Type on Predictions of Solute Mobility in Soil. *Water Resources Research,* 30:3013-3021.

Hjelmar, O., and J. B. Hansen. 2004. Towards Final Storage Quality in Landfilling: An Example. *in* Proceedings of the Waste 2004 Conference 28-30 September 2004, Stratford-Upon-Avon.

Hoen, E., C. A. Johnson, P. Huggenberger, A. Amirbahman, A. Peter, and H. R. Zweifel. 2000. Investigative Strategies and Risk Assessment of Old Unlined Municipal Solid Waste Landfills. *Waste Management Research,* 18:577-589.

Holmes. 1983. The Adsorptive Capacity of Domestic Refuse from Full Scale Active Landfill. *Waste Management,* 73:581-593.

Hoopes, J. A., and D. R. F. Harleman. 1965. *Waste Water Recharge and Dispersion in Porous Media.* Tech. Rep. No 75, Hydrodynamic lab, MIT, Cambridge, Massachusetts.

Huang, K., M. T. vanGenuchten, and R. Zhang. 1996. Exact Solutions for One-Dimensional Transport with Asymptotic Scale-Dependent Dispersion. *App. Math. Modelling,* 20:298-308.

Hudson, A. P., R. P. Beaven, and W. Powrie. 1999. Measurement of the Hydraulic Conductivity of Household Waste in a Large Scale Compression Cell. Pages 461-468 *in* Sardinia 99, Seventh International Waste Management and Landfill Symposium, S Margherita di Pula, Cagliari, Italy.

Hudson, A. P., R. P. Beaven, and W. Powrie. 2001. Interaction of Water and Gas in Saturated Household Waste in a Large Scale Compression Cell. *in* Sardinia 2001, Eighth International Waste Management and Landfill Symposium, S Margherita di Pula, Cagliari, Italy.

Hudson, A. P., R. P. Beaven, J. K. White, and W. Powrie. 2004. Modelling the Compression Behaviour of Landfilled Domestic Waste. *Waste Management,* 24:259-269.

Hunt, A. G. 2005. Percolation Theory and the Future of Hydrogeology. *Hydrogeology Journal,* 13:202-205.

Jaekel, U., A. Georgescu, and H. Vereecken. 1996. Asymptotic Analysis of Nonlinear Equilibrium Solute Transport in Porous Media. *Water Resources Research,* 32:3093-3098.

Jain, P., J. Powell, T. G. Townsend, and D. R. Reinhart. 2006. Estimating the Hydraulic Conductivity of Landfilled Municipal Solid Waste Using the Borehole Permeameter Test. *Journal of Environmental Engineering,* 132:645-652.

Javaux, M., and M. Vanclooster. 2003. Robust Estimation of the Generalised Solute Transfer Function Parameters. *Soil Sci. Am. J.,* 67:81-91.

Javis, N. 1991. *MACRO - a Model of Water Movement and Solute Transport in Macroporous Soils*. Reports and Dissertations. Deptartment of Soil Science, Swedish University Agric. Sci. Uppsala.

Jing-Lin. 2005. *Investigation of Sorption of Tracers in Municipal Solid Waste*. MSc. UCL, London.

Johnson, C. A., G. A. Richner, T. Vitvar, S. N., and M. Eberhard. 1998. Hydrological and Geochemical Factors Affecting Leachate Composition in Municipal Solid Waste Incinerator Bottom Ash. Part 1: The Hydrology of Landfill Lostorf, Switzerland. *Journal of Contaminant Hydrology,* 33:361-376.

Johnston, P., T. C. Atkinson, N. Odling, and J. A. Barker. 2003. Tracer Breakthrough, Hydraulic Properties and Fracture Networks at Sub-Continuum Scales in Aquifers. *in* Conference on Groundwater in Fractured Rocks, Proceedings of the International Association of Hydrogeologists Congress., Prague, Czech Republic.

Johnston, P. B., T. C. Atkinson, N. E. Odling, and J. A. Barker. 2005. Models of Tracer Breakthrough and Permeability in Simple Fractured Porous Media. in Shaw, editor. *Understanding the Micro to Macro Behaviour of Rock - Fluid Systems* Pages 91-102. Geological Society of London.

Jørgensen, P. R., L. D. McKay, and N. H. Spliid. 1998. Evaluation of Chloride and Pesticide Transport in a Fractured Clayey Till Using Large Undisturbed Columns and Numerical Modelling. *Water Resources Research,* 34:539-553.

Jourde, H., F. Cornaton, S. Pistre, and P. Bidaux. 2002. Flow Behavior in a Dual Fracture Network. *Journal of Hydrology,* 266:99-119.

Jury, W. A., and K. Roth. 1990. *Transfer Functions and Solute Transport through Soil: Theory and Applications.* Birkhaeuser Publ. Basel. 235 p.

Kazemi, H. 1969. Pressure Transient Analysis of Naturally Fractured Reservoirs with Uniform Fracture Distribution. *SPE Journal,* Trans AIME Vol 246:451-462.

Kazemi, H., J. R. Gilman, and A. M. El-Sharkawy. 1992. Analytical and Numerical Solution of Oil Recovery from Fractured Reservoirs with Empirical Transfer Functions. *SPE Reservoir Engineering,* SPE19849.

Kazimoglu, Y. K., J. R. McDougall, and I. C. Pyrah. 2006. Unsaturated Hydraulic Conductivity of Landfilled Waste. *in* G. A. Miller, C. E. Zapata, S. L. Houston, and D. G. Fredlund, editors. Unsaturated Soils 2006. Proceedings of the Fourth International Conference on Unsaturated Soils, Carefree, Arizona.

Kinniburgh, D. G. 1986. General Purpose Adsorption Isotherms. *Environmental Science and Technology,* 20:895-904.

Kinniburgh, D. G., J. A. Barker, and M. Whitfield. 1983. A Comparison of Some Simple Adsorption Isotherms for Describing Divalent Cation Adsorption by Ferrihydrite. *Journal of Colloid and Interface Science,* 95:370-384.

Kitanidis, P. K. 1994. The Concept of the Dilution Index. *Water Resources Research,* 30:2011-2026.

Kleisson, F. M., M. B. Beck, and H. S. Wheater. 1990. The Identifiability of Conceptual Hydrochemical Models. *Water Resources Research,* 26:2979-2992.

Klotz, D., and H. Moser. 1974. Hydrodynamic Dispersion as Aquifer Characteristic. in *Isotope Techniques in Groundwater Hydrology*. IAEA, Vienna.

Konikow, L. F., and J. D. Bredehoeft. 1992. Groundwater Models Cannot Be Validated. *Advances in Water Resources,* 15:75-83.

Krause, P., D. P. Boyle, and F. Bäse. 2005. Comparison of Different Efficiency Criteria for Hydrological Model Assessment. *Advances in Geosciences,* 5:89-97.

Kreft, A., and A. Zuber. 1978. On the Physical Meaning of the Dispersion Equation and Its Solutions for Different Initial and Boundary Conditions. *Chem. Eng. Sci.,* 33:1471-1480.

Krueupelbeck, I., and H.-J. Ehrig. 1999. Long-Term Behaviour of Municipal Solid Waste Landfills in Germany. *in* Stegmann and Cossu, editors. Sardinia, S. Margherita di Pula, Cagliari, Italy.

Lapidus, L., and N. R. Amundson. 1952. Mathematics of Adsorption in Beds. IV. The Effects of Longitudinal Diffusion in Ion Exchange and Chromatographic Columns. *Journal of Physical Chemistry,* 56:984-988.

Legates, D. R., and G. J. McCabe. 1999. Evaluating the Use of "Goodness of Fit" Measures in Hydrologic and Hydroclimatic Model Validation. *Water Resources Research,* 35:233-241.

Li, L., D. A. Barry, P. J. Culligan-Hensley, and K. Bajracharya. 1994. Mass Transfer in Soils with Local Stratification of Hydraulic Conductivity. *Water Resources Research,* 30:2891-2900.

Lichtner, P. C. 2000. Critique of Dual Continuum Formulations of Multicomponent Reactive Transport in Fractured Porous Media. in Faybishenko, Witherspoon, and Benson, editors. *Dynamics of Fluids in Fractured Rock.* American Geophysical Union, Washinton DC.

Linhart, H., and W. Zucchini. 1986. *Model Selection.* John Wiley and Sons, Chichester.

Little, R., E. Muller, and R. McKay. 1996. Modelling of Contaminant Migration in a Chalk Aquifer. *Journal of Hydrology,* 175:473-509.

Liu, H. H., G. S. Bodvarsson, and G. Zhang. 2004. Scale Dependency of the Effective Matrix Diffusion Coefficient. *Vadose Zone Journal,* 3:312-315.

Lu, C. 1996. A Model of Leaching Behaviour from MSW Incinerator Residue Landfills. *Waste Management & Research,* 14:51-70.

Ludwig, C., C. A. Johnson, M. Käppeli, A. Ulrich, and S. Riediker. 2000. Hydrological and Geochemical Factors Controlling the Leaching of Cemented MSWI Air Pollution Control Residues: A Lysimeter Field Study. *Journal of Contaminant Hydrology,* 42:253-272.

Ma, L., and H. M. Selim. 1995. Transport of a Nonreactive Solute in Soils: A Two-Flow Domain Approach. *Soil Science,* 159:224-234.

Mackenzie, R., and J. A. Barker. 1998. Mathematical Characterisation of Block and Channel Geometry in Double Porosity Models. *in* Mass Transport in Fractured Aquifers and Aquitards, University of Copenhagen, Denmark.

Maloszewski, P., and A. Zuber. 1985. On the Theory of Tracer Experiments in Fissured Rocks with a Porous Matrix. *Journal of Hydrology,* 79:333-358.

Maloszewski, P., Harum, and Benischke. 1992. Mathematical Modelling of Tracer Experiments in the Karst of the Lurbach System. in Behrens, editor. *Investigations with Natural and Artificial Tracers in the Karst Aquifer of the Lurbach System* Pages 116-143, Graz.

Maloszewski, P., H. Moser, W. Stichler, and P. Trimborn. 1995. Isotope Hydrology Investigations in Large Refuse Lysimeters. *Journal of Hydrology,* 167:149-166.

Maloszewski, P., and A. Zuber. 1990. Mathematical Modeling of Tracer Behavior in Short-Term Experiments in Fissured Rocks. *Water Resources Research,* 26:1517-1528.

Maloszewski, P., and A. Zuber. 1992. On the Calibration and Validation of Mathematical Models for the Interpretation of Tracer Experiments in Groundwater. *Advances in Water Resources,* 15:47-62.

Maloszewski, P., and A. Zuber. 1993. Principles and Practice of Calibration and Validation of Mathematical Models for the Interpretation of Environmental Tracer Data in Aquifers. *Advances in Water Resources,* 16:173-190.

Mantovan, P., and E. Todini. 2006. Hydrological Forecasting Uncertainty Assessment: Incoherence of the GLUE Methodology. *Journal of Hydrology,* 330:368-381.

Margolin, G., B. Berkowitz, and H. Scher. 1998. Structure, Flow, and Generalized Conductivity Scaling in Fracture Networks. *Water Resources Research,* 34:2103-2121.

Martys, N. S., and H. Chen. 1996. Simulation of Multicomponent Fluids in Complex Three-Dimensional Geometries by the Lattice Boltzmann Method. *Physical Review E,* 53:743-750.

Matheron, G., and G. De Marsily. 1980. Is Transport in Porous Media Always Diffusive? A Counterexample. *Water Resources Research,* 16:901-917.

Mathias, S. A. 2005. *Modelling Flow and Transport in the Chalk Unsaturated Zone.* PhD. University of London.

Mathias, S. A., and R. W. Zimmerman. 2003. Laplace Transform Inversion for Late-Time Behaviour of Groundwater Flow Problems. *Water Resources Research,* 39:1283.

Massey, B. S. 1989. *Mechanics of Fluids.* Chapman and Hall, London.

Maurice, L., T. C. Atkinson, J. A. Barker, J. Bloomfield, A. Farrant, and A. Williams. 2004. Karst Development in the Chalk of the Pang Langbourn Catchments. *in* Joint meeting of

the Environmental Geophysics and the Hydrogeology Groups of the Geological Society, The Geological Society, Burlington House, London.
McCreanor, P. T., and D. R. Reinhart. 2000. Mathematical Modeling of Leachate Routing in a Leachate Reciculating Landfill. *Water Research,* 34:1285-1295.
McDougall, J. R., I. C. Pyrah, and S. T. S. Yuen. 2004. Extended Phase Relations and Load Effects in MSW. *Waste Management,* 24:251-257.
McDougall, J. R., and R. S. Silver. 2005. Hydro-Bio-Mechanical Modelling of Landfilled Waste: Real Insights? *in* Cossu and Stegmann, editors. Sardinia 2005, 10th International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Sardinia, Italy.
Meier, P., A. Medina, and J. Carrera. 2001. Geostatistical Inversion of Cross-Hole
Pumping Tests for Identifying Preferential Flow Channels within a Shear Zone. *Ground water,* 39:10-17.
Moench, A. F. 1984. Double-Porosity Models for a Fissured Groundwater Reservoir with Fracture Skin. *Water Resources Research,* 20:831-846.
Moench, A. F. 1995. Convergent Radial Dispersion in a Double-Porosity Aquifer with Fracture Skin: Analytical Solution and Application to a Field Experiment in Fractured Chalk. *Water Resources Research,* 31:1823-1835.
Moreno, A., T. C. Atkinson, and R. Simons. 2003. Fracture Networks: A Case Study from the Chalk. *in* Kasny, Hrkal, and Bruthans, editors. Groundwater in Fractured Rock. Krasny, Prague, Czech Republic: Charles University.
Moreno, L., B. Gylling, and I. Neretnieks. 1997. Solute Transport in Fractured Media - the Important Mechanisms for Performance Assessment. *Journal of Contaminant Hydrology,* 25:283-298.
Moreno, L., and I. Neretnieks. 1985. Analysis of Some Laboratory Tracer Runs in Natural Fissures. *Water Resources Research,* 21:951-958.
Moreno, L., and I. Neretnieks. 1992. Fluid and Solute Transport in a Network of Channels. Pages 691-698 *in* C. G. Sombret, editor. Scientific Basis for Nuclear Waste Management XV. Materials Research Society Symposium Proceedings, Materials Research Society, Pittsburgh, Pennsylvania.
Moreno, L., and C.-F. Tsang. 1994. Flow Channeling in Strongly Heterogeneous Porous Media: A Numerical Study. *Water Resources Research,* 30:1421-1430.
Morris, S. P., and S. P. Watson. 1999. *NAMMU (Release 6.4) Technical Overview.* AEAT-2907 Issue 1, AEA Technology, Harwell, Oxon.
Neretnieks, I. 2002. A Stochastic Multi-Channel Model for Solute Transport - Analysis of Tracer Tests in Fractured Rock. *Journal of Contaminant Hydrology,* 55:175-211.
Neretnieks, I., and A. Ramuson. 1984. An Approach to Modelling Radionuclide Migration in a Medium with Strongly Varying Velocity and Block Sizes Along the Path. *Water Resources Research,* 20:1823-1836.
Nkedi-Kizza, P., J. W. Biggar, M. T. van Genuchten, P. J. Wierenga, H. M. Selim, D. J. M., and D. R. Nielsen. 1983. Modeling Tritium and Chloride 36 Transport through an Aggregated Oxisol. *Water Resources Research,* 19:691-700.
Ogata, A., and R. B. Banks. 1961. *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media.* US Geological Survey Professional Paper 411-A.
Ohlsson, Y., and I. Neretnieks. 1995. *Literature Survey of Matrix Diffusion Theory and of Experiments and Data Including Natural Analogues.* Department of Chemical Engineering and Technology, Royal Institute of Technology.
Olsson, O. 1994. The Site Characterisation and Validation Project. *in* Proc. of the 4th international NEA/SKB Symposium, In situ experiments at the Stripa Mine. OECD, Stockholm.
Öman, C., and H. Rosqvist. 1999. Transport Fate of Organic Compounds with Water through Landfills. *Water Research,* 33:2247-2254.
Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science,* 263:641-646.
Pappenberger, F., and K. J. Beven. 2006. Ignorance Is Bliss: Or Seven Reasons Not to Use Uncertainty Analysis. *Water Resources Research,* 42.

Parker, J. C., and A. J. Valocci. 1986. Constraints on the Validity of Equilibrium and First-Order Kinetic Transport Models in Structured Soils. *Water Resources Research,* 22:399-407.

Parker, J. C., and M. T. vanGenuchten. 1984. Flux-Averaged and Volume-Averaged Concentrations in Continuum Approaches to Solute Transport. *Water Resources Research,* 20:866-872.

Parlange, J.-Y., J. L. Starr, M. T. vanGenuchten, D. A. Barry, and J. C. Parker. 1992. Exit Condition for Miscible Displacement Experiments. *Soil Science,* 153:165-171.

Pavone, and Schwerer. 1983. *Develpment of Coal Gas Production Simulators and Mathematical Models for Well Test Strategies.* Gas Research Institute.

Pedit, J. A., and C. T. Miller. 1994. Heterogeneous Sorption Processes in Subsurface Systems. 1. Model Formulations and Applications. *Environmental Science and Technology,* 28:2094-2104.

Perkins, T. K., and O. C. Johnson. 1963. A Review of Diffusion and Dispersion in Porous Media. *Society of Petroleum Engineers Journal,* 3:70-84.

Pickens, J. F., and G. E. Grisak. 1981. Modeling of Scale-Dependent Dispersion in Hydrogeologic Systems. *Water Resources Research,* 17:1701-1711.

Popper, K. 1979. *Objective Knowledge.* Cambridge University Press, New York.

Poteri, A., D. Bilaux, W. Dershowitz, J. J. Gomez-Hernandez, V. Cvetkovic, A. Hautojarvi, D. Holton, A. Medina, and A. Winberg. 2002. *Final Report of the True Block Scale Project.* TR-02-15, Svensk Kärnbränslehantering AB.

Powrie, W., and J. White. 2004. Landfill Process Modelling. *Waste Management,* 24.

Premchitt, J. 1981. A Technique in Using Integrodifferential Equations for Model Simulation of Multiaquifer Systems. *Water Resources Research,* 17:162-168.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes.* Cambridge University Press, Cambridge.

Pruess, K., and T. N. Narasimhan. 1982. On Fluid Reserves and the Production of Superheated Steam from Fractured, Vapour-Dominated Geothermal Reservoirs. *Journal of Geophysical Research,* 87:9329-9339.

Pruess, K., and T. N. Narasimhan. 1985. A Practical Method for Modeling Fluid and Heat Flow in Fractured Porous Media. *SPE Journal*:14-26.

Pruess, K., and Y. S. Wu. 1993. A New Semi-Analytical Method for Numerical Simulation of Fluid and Heat Flow in Fractured Reservoirs. *SPE Advanced Technology Series,* 1:63-72.

Rabideau, A. J., and C. T. Miller. 1994. Two-Dimensional Modeling of Aquifer Remediaion Influenced by Sorption Nonequilibrium and Hydraulic Conductivity Heterogeneity. *Water Resources Research,* 30:1457-1470.

Rameson, A. 1985. The Effect of Particles of Variable Size, Shape and Properties on the Dynamics of Fixed Beds. *Chemical Engineering Science,* 40:621-629.

Raven, K. G., K. S. Novakowski, and P. A. Lapcevic. 1988. Interpretation of Field Tracer Tests of a Single Fracture Using a Transient Solute Storage Model. *Water Resources Research,* 24:2019-2032.

Redner, S. 2001. *A Guide to First Passage Processes.* Cambridge University Press, Cambridge.

Reedy, O. C., P. M. Jardine, G. V. Wilson, and H. M. Selim. 1996. Quantifying the Diffusive Mass Transfer of Non Reactive Solutes in Columns of Fractured Saprolite. *Soil Sci. Soc. Am. J.,* 60:1376-1384.

Rehfeldt, K. R., J. M. Boggs, and L. W. Gelhar. 1992. Field Study of Dispersion in a Heterogeneous Aquifer. 3. Geostatistical Analysis of Hydraulic Conductivity. *Water Resources Research,* 28:3309-3324.

Reitzel, S., G. Farquahar, and E. McBean. 1992. Temporal Characterisation of Municipal Solid Waste Leachate. *Canadian Journal of Civil Engineering*:668-679.

Rhén, I., G. Gustafson, and P. Wikberk. 1997. *Äspö HRL - Geoscientific Evaluation 1997/4. Results from Pre-Investigations and Detailed Site Characterisation. Comparison of Predictions and Observations Hydrogeology, Groundwater Chemistry and Transport.* TR-97-05, Svensk Kärnbränslehantering AB (SKB), Stockholm.

Robinson, H. D. 1995. *A Review of the Composition of Leachates from Domestic Wastes in Landfill Sites.* DE0918A/FR2, UK Department of Enviroment.

Robinson, H. D., K. Knox, and S. D. Bone. 2004. *Improved Definition of Leachate Source Term for Landfills. Phase 1: Review of Data for European Landfills*. Environment Agency Science Report P1 - 494/SR1, Environment Agency.

Robinson, P. C. 1984. *Connectivity, Flow and Transport in Network Models of Fractured Media*. PhD. Oxford.

Rosqvist, H., and D. Bendz. 1999. An Experimental Evaluation of the Solute Transport Volume in Biodegraded Municipal Solid Waste. *Hydrology and Earth System Science,* 3:429-438.

Rosqvist, H., D. Bendz, R. P. Beaven, and A. Hudson. 2004. Controlled Tracer Tests through Solid Waste in Two Large-Scale Experimental Set-Ups. Pages B3-B14 *in* Nordic Geotechnical Meeting. Swedish Geotechnical Institute, Ystad.

Rosqvist, N. H., L. H. Dollar, and A. B. Fourie. 2005. Preferential Flow in Municpal Solid Waste and Implications for Long-Term Leachate Quality: Valuation of Laboratory-Scale Experiments. *Waste Management Research,* 23:367-380.

Ruthven, D. M., and K. F. Loughlin. 1971. The Effect of Crystallite Shape and Size Distribution on Diffusion Measurements in Molecular Sieves. *Chem. Eng. Sci.,* 26:577-584.

Roux, S., F. Plouraboué, and J.-P. Hulin. 1998. Tracer Dispersion in Rough Open Cracks. *Transport in Porous Media,* 32:97-116.

Sabbas, T., A. Polettini, R. Pomi, T. Astrup, O. Hjelmar, P. Mostbauer, G. Cappai, G. Magel, S. Salhofer, C. Speiser, S. Heuss-Assbichler, and P. Lechner. 2003. Management of Municipal Solid Waste Incineration Residues. *Waste Management,* 23:61-88.

Sagar, B., and A. Runchal. 1982. Permeability of Fracture Rock: Effect of Fracture Size and Data Uncertainty. *Water Resources Research,* 18:266-274.

Sahimi, M. 1994. *Applications of Percolation Theory*. Taylor and Frances, London.

Saltelli, A., K. Chan, and M. Scott. 2000. *Sensitivity Analysis*. John Wiley and Sons, Chichester.

Sanchez-Vila, X., and J. Carrera. 2004. On the Striking Similarity between the Moments of Breakthrough for a Heterogeneous Medium and a Homogeneous Medium with a Matrix Diffusion Term. *Journal of Hydrology,* 294:164-175.

Sardin, M., D. Schweich, F. J. Leij, and M. T. van Genuchten. 1991. Modeling the Nonequilibrium Transport of Linearly Interacting Solutes in Porous Media: A Review. *Water Resources Research,* 27:2287-2307.

Sauty, J.-P. 1980. An Analysis of Hydrodispersive Transfer in Aquifers. *Water Resources Research,* 16:145-158.

Sawada, A., M. Uchida, M. Shimo, H. Yamamoto, H. Takahara, and T. W. Doe. 2000. Non-Sorbing Tracer Migration Experiments in Fractured Rock at the Kamaishi Mine, Northeast Japan. *Engineering Geology,* 56:75-96.

Scheibe, T., and S. Yabusaki. 1998. Scaling of Flow and Transport Behavior in Heterogeneous Groundwater Systems. *Advances in Water Resources,* 22:223-238.

Schumm, S. A. 1991. *To Interpret the Earth. Ten Ways to Be Wrong*. Cambridge University Press, Cambridge.

Schwarz, G. 1978. Estimating the Dimension of a Model. *Annals of Statistics,* 6:461-464.

Selroos, J.-O., D. D. Walker, A. Strom, B. Gylling, and S. Folling. 2002. Comparison of Alternative Modelling Approaches for Groundwater Flow in Fractured Rock. *Journal of Hydrology,* 257:174-188.

Shapiro, A. M. 2001. Effective Matrix Diffusion in Kilometer-Scale Transport in Fractured Crystalline Rock. *Water Resources Research,* 37:507-522.

Šimunek, J., M. T. van Genuchten, and M. Sejna. 2005. *The Hydrus-1D Software Package for Simulating the Movement of Water, Heat, and Multiple Solutes in Variably Saturated Media. Version 3.0, Hydrus Software Series 1*. Department of Environmental Sciences, University of California Riverside, Riverside, California, USA.

Skopp, J., and A. W. Warrick. 1974. A Two-Phase Model for the Miscible Displacement of Reactive Solutes in Soil. *Soil Sci. Soc. Am. Proc.,* 38:545-550.

Spear, R. C., and G. M. Hornberger. 1980. Eutrophication in Peel Inlet, II Identification of Critical Uncertainties Via Generalised Sensitivity Analysis. *Water Research,* 14:43-49.

Spiegel, M. R. 1968. *Mathematical Handbook of Formulas and Tables*. McGraw-Hill, New York.

Srinivasan, P., and J. W. Mercer. 1987. *BIO1D Manual, Version 1.0*. GeoTrans Inc.

Srinivasan, P., and J. W. Mercer. 1988. Simulation of Biodegradation and Sorption Processes in Ground Water. *Groundwater,* 26:475-487.

Stegemann, Lin, and Beaven. 2006. Preliminary Investigation of the Sorption of Common Landfill Tracers to MSW. *in* Waste 2006, Stratford-Upon-Avon.

Stegemann, J., Schneider, B. B. W., and Murphy. 1995. Lysimeter Washing of MSW Incinerator Bottom Ash. *Waste Management & Research,* 13:149-165.

Stegmann, and Ehrig. 1989. Leachate Production and Quality: Results of Landfill Processes and Operation. Proceedings. *in* Cossu, Christensen, and Stegmann, editors. Sardinia 89, Second International Landfill Symposium, S Margherita di Pula, Cagliari, Sardinia, Italy.Straub, W. A., and D. R. Lynch. 1982. Models of Landfill Leaching: Moisture Flow and Inorganic Strength. *Journal of Environmental Engineering, ASCE,* 108:231-250.

Straub, W. A., and D. R. Lynch. 1982. Models of Landfill Leaching: Moisture Flow and Inorganic Strength. *Journal of Environmental Engineering, ASCE,* 108:231-250.

Sudicky, E., and E. Frind. 1982. Contaminant Transport in Fractured Porous Media: Analytical Solutions for a System of Parallel Fractures. *Water Resources Research,* 18:1634-1642.

Sudicky, E. A. 1990. The Laplace Transform Galerkin Technique for Efficient Time-Continuous Solution of Solute Transport in Double-Porosity Media. *Geoderma,* 46:209-232.

Sudicky, E. A., and R. G. McLaren. 1992. The Laplace Transform Galerkin Technique for Large-Scale Simulation of Mass Transport in Discretely-Fractured Porous Formations. *Water Resources Research,* 28:499-514.

Talbot, A. 1979. The Accurate Numerical Inversion of Laplace Transforms. *J. Inst. Math. Appl,* 23:97-120.

Taylor, G. I. 1953. Dispersion of Soluble Matter in Solvent Flowing Slowly through a Tube. *Proc. Royal Soc. A,* 1137:186-203.

Thiemann, M., M. Trosset, H. V. Gupta, and S. Sorooshian. 2001. Bayesian Recursive Parameter Estimation for Hydrologic Models. *Water Resources Research,* 37:2521-2535.

Toride, N., and F. J. Leij. 1996a. Convective-Dispersive Stream Tube Model for Field-Scale Solute Transport: I. Moment Analysis. *Soil Sci. Am. J.,* 60:342-352.

Toride, N., and F. J. Leij. 1996b. Convective-Dispersive Stream-Tube Model for Field-Scale Solute Transport: II. Examples and Calibration. *Soil Sci Am J,* 60:352-361.

Toride, N., F. J. Leij, and M. T. Van Genuchten. 1993. A Comprehensive Set of Analytical Solutions for Nonequilibrium Solute Transport with First-Order Decay and Zero-Order Production. *Water Resources Research,* 29:2167-2182.

Tsang, C. F., and I. Neretnieks. 1998. Flow Channelling in Heterogeneous Fractured Rocks. *Rev. Geophys.,* 36:275-298.

Tsang, Y. W. 1995. Study of Alternative Tracer Tests in Characterising Transport in Fractured Rocks. *Geophys. Res. Lett.,* 22:1421-1424.

Tsang, Y. W., T. C. F., F. V. Hale, and B. Dverstorp. 1987. Tracer Transport in a Stochastic Continuum Model of Fractured Media. *Water Resources Research,* 32:3277-3092.

Tsang, C. F., Y. W. Tsang, and F. V. Hale. 1991. Tracer Transport in Fractures: Analysis of Field Data Based on a Variable-Aperture Channel Model. *Water Resources Research,* 27:3095-3106.

Tsang, Y. W., C. F. Tsang, and F. V. Hale. 1996. Tracer Transport in a Stochastic Continuum Model of Fractured Media. *Water Resources Research,* 32:3077-3092.

Uguccioni, M., and C. Zeiss. 1997. Comparison of Two Approaches to Modelling Moisture Movement through Municipal Solid Waste. *Journal of Environmental Systems,* 23:247-270.

Valocci, A. J. 1985. Validity of the Local Equilibrium Assumption for Modeling Sorbing Solute Transport through Homogeneous Soils. *Water Resources Research,* 21:808-820.

Valocci, A. J. 1990. Use of Temporal Moment Analysis to Study Reactive Solute Transport in Aggregated Porous Media. *Geoderma,* 46:233-247.

van den Daele, G. 2005. *Investigation of Flow and Solute Transport through Unsaturated Chalk.* University of London, London.

van der Sloot, H. A. 2004. Readily Accessible Data and an Integrated Approach Is Needed for Evaluating Waste Treatment Options and Preparation of Materials for Beneficial Use. *Waste Management,* 24:751-752.

van der Sloot, H. A. 2005. End of Black Box Approach? A Step Towards More Sustainable Landfills. *Waste Management,* 25:461.

van der Sloot, H. A., A. van Zomeren, J. J. Dijkstra, J. C. L. Meeuseen, R. N. J. Comans, and H. Scharff. 2005. Prediction of the Leaching Behaviour of Waste Mixtures by Chemical Speciation Modelling Based on a Limited Set of Key Parameters. *in* R. Cossu and R. Stegmann, editors. Sardinia 2005, S. Margherita di Pula, Cagliari, Sardinia, Italy.

van der Zee, S. E. A. T. M. 1990. Analytical Traveling Wave Solutions for Transport with Nonlinear and Nonequlibrium Adsorption. *Water Resources Research,* 26:2563-2578.

van Genuchten, M. T. 1999. *The CXTFIT Code for Estimating Transport Parameters from Laboratory of Field Tracer Experiments.* 137 v.2.1, U.S .Salinity Lab, Riverside, California.

van Genuchten, M. T., and W. J. Alves. 1982. *Analytical Solutions of the One-Dimensional Convective-Dispersive Solute Transport Equation.* Technical Services Bulletin 1661, United States Department of Agriculture.

van Genuchten, M. T., and F. N. Dalton. 1986. Models for Simulating Salt Movement in Aggregated Field Soils. *Geoderma,* 38:165-183.

van Genutchten, M. T. 1980. A Closed-Form Equation for Predicting the Hydraulic-Conductivity of Unsaturated Soils. *Soil Sci. Soc. Am. J.,* 44:892-898.

van Praagh, M., M. Persson, and K. M. Persson. 2005. Assessment of Hydrological Parameters in Column Leaching Tests by Time-Domain Reflectometry. Pages 611-612 *in* R. Cossu and R. Stegmann, editors. Sardinia 2005, Tenth International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Italy.

Visual-Numerics. 1997. *IMSL Fortran Subroutines for Mathematical Applications, Math Library.* Visual-Numerics, Inc, Houston, Texas.

Voss, C. 1984. *SUTRA. A Finite Element Simulation Model for Saturated-Unsaturated, Fluid-Density Dependent Groundwater Flow with Energy Transport or Chemically Reactive Single-Species Solute Transport.* Water Resources Investigations Report 84-4369, US Geological Survey.

Wainwright, J., and M. Mullighan. 2004. *Environmental Modelling. Finding Simplicity in Complexity.* John Wiley & Sons, Ltd.

Wall, D. K., and C. Zeiss. 1995. Municipal Landfill Biodegradation and Settlement. *Journal of Environmental Engineering,* 121:214-224.

Ward, R. S., J. A. Barker, A. T. Williams, L. J. Brewerton, and I. N. Gale. 1998. *Groundwater Tracer Tests: A Review and Guidelines for Their Use in British Aquifers.* WD/98/19 Hydrogeology Series, BGS.

Warrick, A. W., editor. 2002. Soil Physics Companion. CRC Press, Boca Raton.

Warrick, A. W., J. H. Kichen, and J. L. Thames. 1972. Solutions for Miscible Displacement of Soil Water with Time-Dependent Velocity and Dispersion Coefficients. *Soil Sci. Am. Proc.,* 36:863-867.

Watson, S. 2004. *Measurements and Models of Solute Behaviour in the Dual-Porosity Chalk Aquifer of the UK: The Applicability of Tracer Tests.* PhD. University of London, London.

Weglarçzyk, S. 1998. The Interdependence and Applicability of Some Statistical Quality Measures for Hydrological Models. *Journal of Hydrology,* 206:98-103.

Werth, C. J., J. A. Cunningham, P. V. Roberts, and M. Reinhard. 1997. Effects of Grain-Scale Mass Transfer on the Transport of Volatile Organics through Sediments. *Water Resources Research,* 33:2727-2740.

White, J., J. Robinson, and Q. Ren. 2004. Modelling the Biochemical Degradation of Solid Waste in Landfills. *Waste Management,* 24:227-240.

White, J. K., Q. Ren, and J. P. Robinson. 2003. A Framework to Contain a Spatially Distributed Model of the Degradation of Solid Waste in Landfills. *Waste Management and Research,* 21:330-345.

White, R. E., J. S. Dyson, R. A. Haigh, W. A. Jury, and G. Sposito. 1986. A Transfer Function Model of Solute Transport through Soil 2. Illustrative Applications. *Water Resources Research,* 22:248-254.

Witherspoon, P. A., Wilson, Long, Galbraith, and Dubois. 1980. Measure de Perméabilité en Grand dans les Roches Cristallines Fracturées. *Bull BRGM,* III(1):53-61.

Wood, A. L., and J. M. Davidson. 1975. Fluometuron and Water Content Distributions During Infiltration: Measured and Calculated. *Soil Sci. Am. Proc.,* 39:820-825.

Woodbury, A. D., and Y. Rubin. 2000. A Full-Bayesian Approach to Parameter Inference from Tracer Travel Time Moments and Investigation of Scale Effects at the Cape Cod Experimental Site. *Water Resources Research,* 36:159-171.

Woodman, N. D. 2004. *Macroscopic Mass Transfer Models of Solute Transport in the Chalk.* WS Atkins.

Woodman, N. D., J. A. Barker, and R. P. Beaven. 2005. Identification of Transport Processes from a Tracer Test through Waste. *in* S. Cossu, editor. Sardinia 2005 Tenth International Waste Management and Landfill Symposium, S Margherita di Pula, Cagliari, Sardinia, Italy.

Wright, T. 2002. *Predicting the Applicability of Aquifer Storage Recovery (ASR) to the UK Chalk Aquifer.* University College London, London.

Wright, T. E. J., A. C. Carruthers, and A. J. 2005. *Chalk FTI (Biorem 36): Model Code Review.* 5030710/2001326/R004. 32p.

Wu, Y. S., J. B. Kool, P. S. Huyakorn, and Z. A. Saleem. 1997. An Analytical Model for Nonlinear Adsorptive Transport through Layered Soils. *Water Resources Research,* 33:21-29.

Wu, Y.-S., H. H. Liu, and G. S. Bodvarsson. 2004. A Triple-Continuum Approach for Flow and Transport Processes in Fractured Rock. *Journal of Contaminant Hydrology,* 73:145-179.

Yildiz, E. D., K. Ünlü, and R. K. Rowe. 2004. Modelling Leachate Quality and Quantity in Municipal Solid Landfills. *Waste Management & Research,* 22:78-92.

Young, D. F., and W. P. Ball. 1995. Effects of Column Conditions on the First-Order Rate Modelling of Nonequilibrium Solute Breakthrough. *Water Resources Research,* 31:2181-2192.

Young, D. F., and W. P. Ball. 1997. Effects of Column Conditions on the First-Order Rate Modelling of Nonequilibrium Solute Breakthrough: Cylindrical Macropores Versus Spherical Media. *Water Resources Research,* 33:1149-1159.

Young, P. C., A. Chotai, and K. Beven. 2004. Data-Based Mechanistic Modelling and the Simplification of Environmental Systems. in J. Wainwright and M. Mullighan, editors. *Environmental Modelling. Finding Simplicity in Complexity.* Wiley, Chichester.

Yu, J.-W., and I. Neretnieks. 1997. *Diffusion and Sorption Poperties of Radionuclides in Compacted Bentonite.* 97-12, Svensk Kärnbränslehantering AB, Stockholm.

Yuen, S. T. S., T. A. McMahon, and J. R. Styles. 2000. Monitoring in Situ Moisture Content of Municipal Solid Waste Landfills. *Journal of Environmental Engineering,* 126:1088-1095.

Zacharof, A. I. 2001. *Stochastic Modelling of Landfill Leachate and Biogas Production Incorporating Waste Heterogeneity.* PhD. University of London, London.

Zacharof, A. I., and A. P. Butler. 2004a. Stochastic Modelling of Landfill Leachate and Biogas Production Incorporating Waste Heterogeneity. Model Formulation and Uncertainty Analysis. *Waste Management,* 24:453-462.

Zacharof, A. I., and A. P. Butler. 2004b. Stochastic Modelling of Landfill Processes Incorporating Waste Heterogeneity and Data Uncertainty. *Waste Management,* 24:241-250.

Zhang, R. 2000. Generalised Transfer Function Model for Solute Transport in Heterogeneos Soils. *Soil Sci. Soc. Am. J.,* 64:1595-1602.

Zhang, D., K. Beven, and A. Mermoud. 2006. A Comparison of Non-Linear Least Square and GLUE for Model Calibration and Uncertainty Estimation for Pesticide Transport in Soils. *Advances in Water Resources,* 29:1924-1933.

Zimmerman, R. W., and G. S. Bodvarsson. 1989. An Approximate Solution for One-Dimensional Absorption in Unsaturated Porous Media. *Water Resources Research,* 25:1422-1428.

Zimmerman, R. W., and G. S. Bodvarsson. 1990. Absorption of Water into Porous Blocks of Various Shapes and Sizes. *Water Resources Research,* 26:2797-2806.

Zimmerman, R. W., G. Chen, T. Hadgu, and G. S. Bodvarsson. 1993. A Numerical Dual-Porosity Model with Semianalytical Treatment of Fracture/Matrix Flow. *Water Resources Research,* 29:2127-2137.

Zinn, B., and C. F. Harvey. 2003. When Good Statistical Models of Aquifer Heterogeneity Go Bad: A Comparison of Flow, Dispersion, and Mass Transfer in Connected and Multivariate Guassian Hydraulic Conductivity Fields. *Water Resources Research,* 39:1051, doi:1010.1029/2001WR001146.

A1 APPENDIX 1 – Transport codes

A1.1 Introduction

The most relevant models codes that have been developed to produce or to test the results in the thesis are detailed here and summarised in Table A1.1. They are predominantly analytical or semi-analytical solutions, supplemented by numerical methods where appropriate. The FORTRAN90 codes were compiled in Compaq Digital FORTRAN 6.5 and make extensive use of the IMSL library.

Table A1.1: Summary of codes in Appendix and their capabilities

Code	Language	Author	Description
Primary codes			
DPD-PULSE	FORTRAN77	Modifications to Barker's DP-PULSE code by Woodman	Hydraulic dispersion in the mobile zone is retrospectively added to the DP-PULSE code.
LT-INVERT	FORTRAN90	Woodman	Suite of LT transport codes, including the following capabilities: • Inversion of AD-DP equations • Finite downstream boundary • Convolution with mixer • Mass recovery • Mean immobile concentration • Flushing system
HET-DP	FORTRAN90	Woodman	Heterogeneous BGF (i.e. multiple block sizes with a distribution of characteristic times, $P(t_{cb})$
ARDVARCS	FORTRAN90	Woodman	Semi-analytical solution to the ADE for piecewise constant input $V(0,t)$ and $C_0(0,t)$
BRENNER-CLEARY	VISUAL BASIC for APPLICATIONS (VBA)	Woodman	Time-domain analytical solutions to ADE equation for a type 1 boundary (Cleary and Adrian, 1973) and a type 3 boundary (Brenner, 1962), with a finite zero concentration gradient downstream boundary. (NB supplements and checks the LT solution)
HET-STREAMS	FORTRAN90	Woodman	Heterogeneous streamtubes (i.e. multiple streamtubes with a distribution of advection times, $P(t_a)$)
SHARP-SORB	FORTRAN90	Woodman (based on Wu et al. (1997))	Semi-analytical code for advection and non-linear sorption in the absence of hydraulic dispersion.
SORBER	FORTRAN90	Woodman (based on Wood and Davidson (1975))	Finite difference code for advection-dispersion with nonlinear sorption
MONTY	FORTRAN90	Woodman	Suite of Monte Carlo codes

Code	Language	Author	Description
		(based on Press et al. (1992))	
Ancillary codes			
DOUBLET	Excel spreadsheet	Woodman	Step solution to doublet transport in the absence of AD within streamtubes (Hoopes and Harleman, 1965)
HYPERCONV adaptations	C++	Adaptations of Peter Robinson's code by Woodman	Cylindrical and spherical boundary conditions, histograms of head and number of 'active' and 'flow' nodes as well as a 'classical' network generator
INTERSECT	FORTRAN90	Woodman	Tests an analytical solution for the probability of a node being intersected by a channel in HYPERCONV
ERROR-SPAC	FORTRAN90-MATLAB®	Woodman	Contour plots of error between model and data
CONVOL	FORTRAN90	Woodman	LT-based convolution of known transfer function and a time-variable input function
VAR-VEL	FORTRAN90	Woodman	Variable flow solution of Warrick et al. (1972)
CUM-ADE	Excel	Woodman	Solution of Ogata and Banks (1961) reparameterised in terms of cumulative volume

A1.2 Dual Porosity codes

A1.2.1 DPD-PULSE (FORTRAN 77)

DPD-PULSE is a version of DP-PULSE (Barker et al., 2000), with hydraulic dispersion in the mobile zone. Since the advective component of DP-PULSE is simulated numerically it is natural to simulate hydraulic dispersion numerically. The DP exchange component therefore remains as an analytical solution for each cell.[1]

Advection is simulated in DP-PULSE by solute being moved downstream in discrete pulses. This occurs at a rate of VS cells per unit timestep. Given that there are $NANN$ cells and solute starts in cell 1, the mobile advection time through the whole system is $t_A = (NANN - 1)/VS$. Without any DP component (e.g. if $\sigma = 0$), a pulse of solute would pass through the system in time, t_A, in the manner of undispersed 'piston flow'.

[1] The various publication on DP-PULSE do not give any detailed description of the code itself. The codes FLOW and FLOPPY (available from netlib) were used to assist in producing a pseudocode which assisted the retrospective additions presented here.

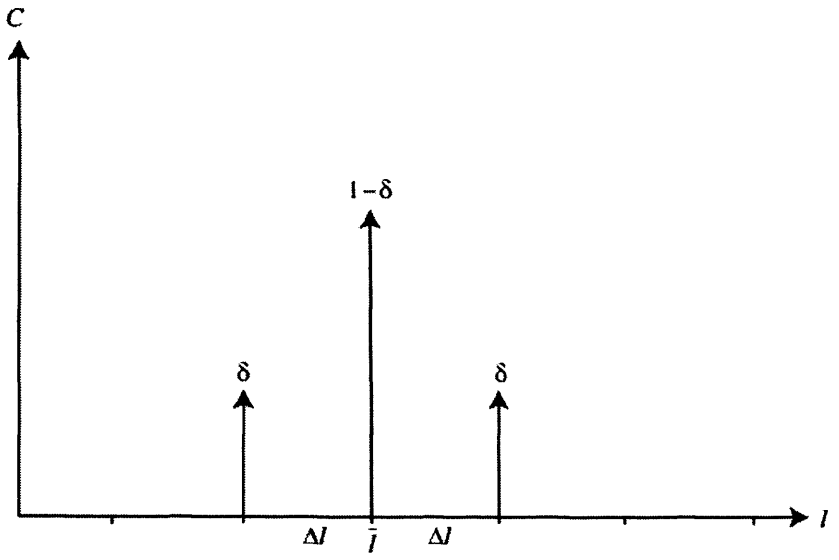


Figure A1.2: Concentrations after one numerical mixing timestep (where in the previous timestep $C = 0$ at every node except for one node, where $C = 1$)

Considering Figure A1.2, dispersion for a mixing timestep is calculated as $\alpha = \frac{\sigma_l^2}{2\bar{l}} = \frac{\delta \Delta l}{2}$, or, $D = \frac{\delta \Delta l^2}{2\Delta t}$. Given $\Delta L = L/(NANN-1)$, the DPD-PULSE dispersion parameter, δ, can be set to achieve the required dispersivity, α. Since $0 \le \delta < 1$, the widest constraint on dispersivity is given by $0 \le \alpha < \frac{\Delta l}{2}$. A better, but narrower constraint is $0 \le \delta < 2/3$ and $0 \le \alpha < \frac{\Delta l}{3}$, which ensures that the numerical dispersion does not temporarily cause the physically impossible situation after a dispersion timestep that a cell with mass in it ends up with less mass than its adjacent cells.

Two key aspects of this method were checked against a FORTRAN90 code (CELLMIXER):

- The accuracy of the predicted dispersion (in comparison to an exact Gaussian)
- The upper limit to dispersion

The accuracy would be expected to be poor for few timesteps, but to improve over a larger number of dispersions. However by 12 steps the accuracy is already high. This is demonstrated spatially in Figure A1.3, for maximum $\delta = 2/3$.

- The code uses this guide value to decide on the number of loops: $N_L = \frac{2\alpha}{\delta_0 \Delta l}$
- δ is then calculated more precisely, as $\delta = \frac{2\alpha}{N_L \Delta l}$

The algorithm, having been tested separately from DP-PULSE, was then incorporated into DP-PUSLE. Testing occurred firstly for a case whereby the DP system was inhibited (i.e. a test of the pure hydraulic dispersion as above, but for a BTC). Secondly the simultaneous effect of AD and DP were tested against analytical solutions (DP1D). Figure A1.4 shows the BTC at the last cell ($NANN$ =51).

Calculations are performed in double precision, in order to enable a wide range of concentrations to be evaluated.

The mass recovery and the immobile concentrations are given here.

1.2.2.2 Mass recovery

A general expression for residual mass $M(t)$ left in a flushing system of length L, at time t, is:

$$M(t) = \underbrace{M_0}_{initial\_mass} \underbrace{- A\theta_m V_m \int_0^t C_m(L,t)dt}_{mass\_out} \qquad \text{A(1.1)}$$

where $M_0 = C_I(\theta_{im} + \theta_m)AL$ is the initial mass in the system, C_I is the initial concentration in the system, $C_m(L,t)$ is the mobile concentration leaving the system at depth L and A is cross sectional area.

Taking LT and rearranging gives the mass fraction remaining:

$$\frac{\overline{M}}{M_0} = \frac{1}{s}\left[1 - \frac{V_m \overline{C}_m}{sC_I(\sigma+1)L}\right] \qquad \text{A(1.2)}$$

Note that this is a dimensioned and corrected version of the erroneous Eq B-6 in Haggerty and Gorelick (1998).

1.2.2.3 Immobile concentration

Given an initially solute-free system, by rearranging the BGF, $B = \frac{\langle \overline{C}_{im} \rangle}{\overline{C}_m}$, the average immobile concentration is $\langle \overline{C}_{im} \rangle = B.\overline{C}_m$. Therefore if $\overline{C}_m$ is known, it is straightforward to compute $\langle \overline{C}_{im} \rangle$.

In the event of locally stochastic heterogeneous blocks described by multiple discrete blocks (see Chapter 7) it is straightforward to prove that the average concentration left in a given block size class is, $\overline{C}_{im1} = B_1.\overline{C}_m$ and $\overline{C}_{im2} = B_2.\overline{C}_m$ and so on. This approach could potentially be extended to a continuous distribution of blocks.

A1.2.3 HET-DP

The LT of the heterogeneous DP system,

$$\bar{C}_f = C_I e \exp\left[-t_a s\left(1+\sigma\int_0^\infty \mathrm{P}(t_{cb})\mathrm{B}(\sqrt{st_{cb}})dt_{cb}\right)\right] \qquad \text{A(1.3)}$$

is sought. In general, this integral is difficult to determine analytically. For computational purposes $\mathrm{B}(s)$ is made in to a truncated infinite sum, and integrated using the midpoint rule. For example for a slab geometry, $\mathrm{B}(s) = \sum_{N=1}^{M} \mathrm{P}(t_{cbi}) \tanh\left(\sqrt{st_{cbi}}\right)\Big/\sqrt{st_{cbi}}$, where M is a large finite number and the steps Δt_{cb} are kept small.

It would be worth experimenting with more sophisticated quadrature in the future.

A1.3 Advection dispersion codes

A1.3.1 ARDVARCS (FORTRAN 90)

There are no open source or commercially available AD codes which offer the ability to input piecewise constant V and C. ADVARCS (Advection-dispersion with Restated Velocity And Restated ConcentrationS) was therefore created.

Of available codes, CXTFIT (van Genuchten, 1999) gets the closest to this need. It is possible to have a piecewise constant C in CXTFIT, but not V at the same time. However the Green's functions tabulated in Table 2.2 of the CXTFIT manual assist towards the more general solution[3]. Before adoption into ARDVARCS the functions in Table 2.2 have been checked in Derive[4] for self-consistent solutions to the ADE and also the relationships between flux-averaged and resident solutions have been tested. They were confirmed as correct, with the exception of a small typo.[5]

ADVARCS solves the basic 1D ADE for resident concentration, , $C_r(t)$ at $z = L$ and then converts this to flux-averaged concentration, $C_f(t)$.

---

[3] The full derivations for the functions in Table 2.2 are found in Toride et al. (1993). The derivation for the initial value problem is repeated in Jury and Roth (1990:186) and is considerably clearer than in Toride et al. (1993) because it includes fewer processes. It is straightforward to show, by dimensionalising the variables and eliminating unused processes, that that the solutions are identical.

[4] Derive 4.41 (Soft Warehouse Inc) is a computer algebra system.

[5] Note Γ_1 for C_f in Table 2.2 of van Genuchten (1999) should be corrected to read:

$$\exp\left[-\mu\tau / R\right]\sqrt{\frac{RPZ^2}{4\pi\tau^3}}\exp\left[-\frac{P(RZ-\tau)^2}{4R\tau}\right]$$

The ADE in terms of $C_r(t)$ is,

$$\frac{\partial C_r}{\partial t} = D\frac{\partial^2 C_r}{\partial z^2} - V\frac{\partial C_r}{\partial z} \qquad \text{A(1.4)}$$

This is solved subject to the following conditions:

Initial condition: $C_r(z,0) = CI_r^0 = 0$ for $\forall x > 0$

Downstream Boundary condition: $\left.\frac{\partial C_r}{\partial z}\right|_{z\to\infty} = 0$

Upstream boundary condition: Both the flow velocity $V(t)$ and the flux-averaged inlet concentration $C_f(0,t)$ are piece-wise constant in time.

So for:

$t = 0$ to t_1: $C_f(0,t) = \mathrm{CB}_f^0$ and $V = V_0$ over a timestep of Δt_1

$t = t_1$ to t_2: $C_f(0,t) = \mathrm{CB}_f^1$ and $V = V_1$ over a timestep of Δt_{N-1}

$t = t_N$ to t_{N+1}: $C_f(0,t) = \mathrm{CB}_f^N$ and $V = V_N$ over a timestep of Δt_N

i.e. the upstream boundary condition is described by constant mass flux:

$$\left. C_r^0 - \frac{D}{V}\frac{\partial C_r^0}{\partial z}\right|_{z=0} = C_f^0(0,t) = \mathrm{CB}_f^0$$

Semi-analytical solution

So, for a given stress period starting at t_N, Δt_N, there is a flow velocity and inlet concentration and the ADE must be solved given the initial distribution of concentration within the system. The concentration at a general time, t, is found by solving the ADE for each stress period, using the distribution of concentration at the end of the last period as the initial condition for the next and so on up until t.

Apart from the first stress period, where $C_r(z,0) = \mathrm{CI}_r^0 = 0$, the initial condition for period N is given by $\mathrm{CI}_r(z,t_N) = \mathrm{CI}_r^N$.

The solution to the ADE is the linear superposition of the solution to the boundary value problem, with $\mathrm{CI}_r = 0$, and the solution to the initial condition problem, with $\mathrm{CI}_f^N = 0$

(Farlow, 1982). The solutions to both parts of the equation (using dimensionless variables $Z = z/L$, $\tau = Vt'/L$, $Pe = VL/D$) are (van Genuchten, 1999):

$$C_r^N(Z,T) = \text{CB+CI} = \underbrace{\int_0 C(0,\tau)\Gamma_B(Z,T-\tau)d\tau}_{\text{CB}} + \underbrace{\int_0^\infty C(\eta,0)\Gamma_I(Z,\eta,\tau)d\eta}_{\text{CI}} \qquad \text{A(1.5)}$$

The first part of the equation for the boundary value problem is a convolution integral, where Γ_B is the impulse response, to a type-3 inlet condition in C_r. The second part of the equation for the initial condition is a spatial integration.[6]

$C_r(Z,T)$ must be solved for each stress period, with $C_r(Z,T)$ at the end of the period becoming $\text{CI}_r(Z)$ for the initial value problem in the next period.

van Genuchten (1999) gives Γ_B and Γ_I as:

$$\Gamma_I = \sqrt{\frac{Pe}{4\pi\tau}}e^{-\frac{Pe(\eta-Z+\tau)^2}{4\tau}} + \sqrt{\frac{Pe}{4\pi\tau}}e^{PeZ}e^{-\frac{Pe(\eta+Z+\tau)^2}{4\tau}} - \frac{Pe}{2}e^{PeZ}\text{erfc}\left(\frac{\eta+Z+\tau}{\sqrt{4\tau/Pe}}\right) \qquad \text{A(1.6)}$$

$$\Gamma_B = \sqrt{\frac{Pe}{\pi\tau}}e^{-\frac{Pe(Z-\tau)^2}{4\tau}} - \frac{Pe}{2}e^{PeZ}\text{erfc}\left(\frac{Z+\tau}{\sqrt{4\tau/Pe}}\right) \qquad \text{A(1.7)}$$

CB can be found without needing to solve the convolution integral, as for a given time period the $C(0,\tau)$ term is a constant. The solution to the integral for CB is simply the type-3 solution[7]:

$$\frac{\text{CB}}{C_0} = \frac{1}{2}\text{erfc}\left(\frac{z-Vt}{\sqrt{4Dt}}\right) + \left(\frac{V^2t}{\pi D}\right)^{1/2}e^{\frac{-(z-Vt)^2}{4Dt}} - \frac{1}{2}\left(1+\frac{Vz}{D}+\frac{V^2t}{D}\right)e^{\frac{Vz}{D}}\text{erfc}\left(\frac{z+Vt}{\sqrt{4Dt}}\right) \qquad \text{A(1.8)}$$

For CI, the integral must be evaluated. This could be performed numerically within FORTRAN, but because this routine must be called for every step in time, it becomes computationally expensive[8]. To speed the code up the spatial integral is firstly truncated, and then split into a summation:

$$C_I^r(Z,\tau) = \int_0^{L_{\max}} C_i(\eta,0)\Gamma_I(Z,\eta,\tau)d\eta \qquad \text{A(1.9)}$$

[6] The derivation of the initial condition solution is shown by Toride et al. (1993), but is very cumbersome as includes decay terms, retardation and first order transfer terms. A cleaner solution is given in Jury and Roth (1990:186). Note that η is dummy Z variable for the integration.

[7] NB: this Equation is quoted incorrectly in the book 'Contaminant Hydrogeology' (Fetter, 1999)

[8] The same code was constructed using the DQDAGS integral in IMSL. This made a useful verification test against the numerical integral, but was considerably slower and therefore is not taken forwards into the final ARDVARCS code.

$$C_I^{\,r}(Z,\tau)=\sum_{i=0}^{N-1}\int_{i\Delta Z}^{(i+1)\Delta Z} C_i(\eta,0)\Gamma_I(Z,\eta,\tau)d\eta \text{ , for } N=L_{\max}/\Delta z \qquad \text{A(1.10)}$$

The integral of a dimensioned version of Equation A(1.10) has been evaluated in Derive, giving a somewhat complicated expression which was exported directly to FORTRAN.

The FORTRAN code was tested for a simple spatial 'top-hat' initial condition, and zero inlet boundary condition.

Initial condition: $\mathrm{CI}_r(z,0)=C_1$, for $0\le z<z_1$ and $\mathrm{CI}_r(z,0)=0$, for $z\ge z_1$

Inlet boundary condition: $\mathrm{CB}_r(0,t)=0$

The solution to this is given in van Genuchten and Alves, 1982, equation (A6):

$$\frac{\mathrm{CI}_r}{C_1}=\left[\begin{array}{l}\dfrac{1}{2}\operatorname{erfc}\left(\dfrac{z-z_1-Vt}{\sqrt{4Dt}}\right)+\left(\dfrac{V^2t}{\pi D}\right)^{1/2}e^{\frac{Vz}{D}}-\dfrac{(z+z_1+Vt)^2}{4Dt}\\ -\dfrac{1}{2}\left(1+\dfrac{V(z+z_1)}{D}+\dfrac{V^2t}{D}\right)e^{\frac{Vz}{D}}\operatorname{erfc}\left(\dfrac{z+Vt}{\sqrt{4Dt}}\right)\end{array}\right]-$$
$$\left[\begin{array}{l}\dfrac{1}{2}\operatorname{erfc}\left(\dfrac{z-Vt}{\sqrt{4Dt}}\right)+\left(\dfrac{V^2t}{\pi D}\right)^{1/2}e^{\frac{-(z-Vt)^2}{4Dt}}\\ -\dfrac{1}{2}\left(1+\dfrac{Vz}{D}+\dfrac{V^2t}{D}\right)e^{\frac{Vz}{D}}\operatorname{erfc}\left(\dfrac{z+Vt}{\sqrt{4Dt}}\right)\end{array}\right] \qquad \text{A(1.11)}$$

The solutions given by ARDVARCS and Equation A(1.11) matched even for extreme values of α and V – thereby verifying the integration.

Finally, to achieve a flux-averaged BTC at a column outlet, the solution to $C_f(L,t)$ is found by, $C_f(L,t)=C_r(L,t)-\dfrac{D}{v}\dfrac{\partial C_r}{\partial z}\Big|_{L,t}$. The calculation of the $\dfrac{\partial C_r}{\partial z}\Big|_{L,t}$ term is performed using numerical central difference for the nodes either side of $L=z(N)$

Hence: $C_f\ (L,t)\ =\ C_r(N)\ -\ \dfrac{C_r(N+1)-C_r(N-1)}{2\Delta z}$

A further version of ARDVARCS has been produced in which numerical mixing has been introduced to allow simulation of up- or downstream mixing (say in a well, or column experiment inlet/outlet system).

A1.3.2 BRENNER-CLEARY (Visual Basic)

Van Genuchten and Alves (1982) gave a wide range of solutions to different inlet and outlet boundary conditions for the ADE.

For $C_r = H(t)$ and $\left.\frac{dC_r}{dz}\right|_{z=L} = 0$, van Genuchten and Alves (1982: Equation A3) cited the solution by Cleary and Adrian (1973):

$$C(z,t) = 1 - \sum_{m-1}^{\infty} \frac{2\beta_m \sin\left(\frac{\beta_m z}{L}\right) \exp\left[\frac{Vz}{2D} - \frac{V^2 t}{4D} - \frac{\beta_m{}^2 Dt}{L^2}\right]}{\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2 + \frac{VL}{2D}} \qquad \text{A(1.12)}$$

where the eigenvalues β_m are the positive roots of:

$$\beta_m \cot(\beta_m) + \frac{vL}{2D} = 0 \qquad \text{A(1.13)}$$

For $C_f = H(t)$ and $\left.\frac{dC_r}{dz}\right|_{z=L} = 0$, van Genuchten and Alves (1982:Equation A3) cited the solution by Brenner (1962):

$$C(z,t) = 1 - \sum_{m-1}^{\infty} \frac{\frac{2VL}{D}\beta_m \left(\beta_m \cos(\frac{\beta_m z}{L}) + \frac{VL}{2D}\sin\left(\frac{\beta_m z}{L}\right)\right) \exp\left[\frac{Vz}{2D} - \frac{V^2 t}{4DR} - \frac{\beta_m{}^2 Dt}{L^2 R}\right]}{\left(\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2 + \frac{VL}{D}\right)\left(\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2\right)} \qquad \text{A (1.14)}$$

where the eigenvalues β_m are the positive roots of:

$$\beta_m \cot(\beta_m) - \frac{\beta_m{}^2 D}{VL} + \frac{VL}{4D} = 0 \qquad \text{A(1.15)}$$

Both Brenner and Cleary's solutions have been evaluated in VBA and can be called from Excel in BrennerCleary.xls. The positive roots are found using the Newton-Raphson method. The impulse responses are easily achieved within the spreadsheet by differentiating these solutions.

A1.3.3 HET-STREAMS (FORTRAN90)

The multi-streamtube model, $\bar{C} = (m_0 / V) \int_0^{\infty} \frac{t_a}{\sqrt{\omega \pi t^3}} e^{\frac{-(t_a - t)^2}{\omega t}} p(t_a) dt_a$ is solved by a discrete sum, either by the midpoint rule: $\bar{C} = (m_0 / V) \Delta t_a \sum_{i=1}^{M} \frac{t_{ai}}{\sqrt{\omega \pi t^3}} e^{\frac{-(t_{ai} - t)^2}{\omega t}} p(t_{ai})$, where $\Delta t_a = t_a^{MAX} / M$, or by the trapezium rule. More advanced forms of quadrature may become necessary in the instance of more demanding $p(t_{ai})$.

A1.4 Sorbtion codes

A1.4.1 SHARP-SORB (FORTRAN90)

The transport equation for non-linear sorption in the absence of hydraulic dispersion (Equation 5.40) is restated here:

$$\left(1 + \frac{B_d}{\theta} \mathrm{f}'(C)\right) \frac{\partial C}{\partial t} = -V \frac{\partial C}{\partial z} \qquad \text{A(1.16)}$$

To model the response to a finite-duration input in the absence of hydraulic dispersion, but in the presence of non-linear sorption, requires a model that will describe a sharp front on one side of the solute pulse and a spreading front on the other (which side will depend on the injection condition and the curvature of the isotherm).

The sharp front tends to cause numerical difficulties for finite difference solutions. Therefore, Wu et al. (1997) described a semi-analytical method for computing such a response and the method has been coded here. The method employs a mass-balance to calculate two fronts on either side of a pulse as it moves through a column, with one of the fronts as a sharp shock and the other as a dispersed front. At the outset the code tests to see whether it is the upstream front or the downstream front that is a shock.

This is a useful tool as it provides at a semi-analytical end-member solution for which more general codes in which $D > 0$ can be tested. Take, for example, the case of $N < 0$ and a finite pulse 'top hat' inlet of tracer. The upstream side of the pulse starts as a sharp shock and remains so. The back of the pulse starts as a sharp step, but rapidly disperses. This dispersion 'eats' away the pulse and eventually reduces the peak concentration.

where subscript i denotes space and superscript j denotes time. DNUM is a correction for numerical dispersion: $DNUM = \frac{V\Delta z}{2} - \frac{V^2\Delta t}{2}$. For Freundlich sorption, $R_i^j = 1 + A.N.C_i^{j^{N-1}}$, where $A = \frac{B_d K_d}{\theta}$.

SORBER has been successfully tested against the solution given in Srinivasan and Mercer, 1987, 1988. SORBER has also been compiled as a DLL and has been linked so can be run from an Excel spreadsheet.

A1.5 Monte Carlo codes – MONTY (FORTRAN 90)

The mathematics and methods employed in MONTY.F90 are entirely standard. However, rather than use an 'off the shelf' Monte Carlo code, a new code has been written. This has the advantage that it is readily adapted to the job in hand. Additionally, the use of automatic fitting and other computationally intensive methods are becoming increasingly standard for groundwater transport applications (Carrera et al., 2005), yet have sufficient pitfalls to require engagement rather than passive-application. Press et al. (1992) writes:

> "Offered the choice between mastery of a five-foot shelf of analytical statistics books and middling ability at performing statistical Monte Carlo simulations, we would surely choose the latter skill"

The most basic assumption that underlies MC relies on the shape of the distribution $\underline{a}_i - \underline{a}_{\text{true}}$ being about the same as $\underline{a}_i - \underline{a}_0$, where $\underline{a}_0$ are the best fit parameters to the original data, $\underline{a}_i$ are the parameters for a given fit and $\underline{a}_{\text{true}}$ are the true parameters. This is reasonable assumption, as long as the way in which random errors occur do not vary rapidly as a function of $\underline{a}_{\text{true}}$.

Press et al., 1992 describes the goodness-of-fit by a model with vector of M parameters '$\underline{a}$'. The outputs C, are compared to data points C_i at discrete points, t_i, giving:

$$\chi^2 = \sum_{i=0}^{N-1} \left(\frac{C_i - C(t_i;\underline{a})}{\sigma_i} \right)^2 \qquad \text{A(1.19)}$$

A Taylor series expansion of χ^2, for variations of about the point of best fit, where $\underline{\delta a} = \underline{a} - \underline{a}_0$ yields a first term which is just a constant. The second term is the gradient which

is zero at the minimum. The third term is $\frac{1}{2}\underline{\delta a}^T \frac{\partial^2 \chi^2}{\partial x_k \partial x_l}\underline{\delta a}$. Defining the second derivative matrix (the 'Hessian') as $\alpha_{kl} = \frac{1}{2}\frac{\partial^2 \chi^2}{\partial x_k \partial x_l} = [\alpha]$, makes the third term: $\underline{\delta a}^T [\alpha] \underline{\delta a}$. It is possible to show that $\Delta\chi^2 = \underline{\delta a}^T [\alpha] \underline{\delta a}$.

For normally-distributed errors, the higher terms in the series are zero. For other error distributions the higher terms are in general not zero, though it may be possible in some instances to approximate them as zero, especially for small $\underline{\delta a}$.

In two dimensions it is clear from this expansion that χ^2 is described by a quadratic, so the loci of constant values of χ^2 will be ellipses. By analogy it is now possible to recognise a multidimensional quadratic form and anticipate multidimensional 'ellipsoids' of constant χ^2 surfaces.

To determine whether points are outside the Confidence Region (CR), all that is required is to determine if $\underline{\delta a}^T [\alpha] \underline{\delta a} - \Delta\chi^2 > 0$. If $\underline{\delta a}^T [\alpha] \underline{\delta a} - \Delta\chi^2 \leq 0$, then the points are on or in the region.

Changing to a coordinate system aligned with the eigenvectors of $[\alpha]$ is achieved by diagonalising $[\alpha]$ by transforming $a = Py$, where P is a matrix whose columns are the eigenvectors of $[\alpha]$. So, $\underline{\delta y}^T \underbrace{P^T \frac{[\alpha]}{\Delta\chi^2} P}_{D} \underline{\delta y} = 1$, where D is a diagonal matrix containing the eigenvalues, $D = \frac{1}{\Delta\chi^2}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. Multiplying out $\underline{\delta y}^T D \underline{\delta y} = 1$ gives, $\frac{y_1^2}{\left(\Delta\chi^2/\lambda_1\right)} + \frac{y_2^2}{\left(\Delta\chi^2/\lambda_2\right)} = 1$, which is the equation of an ellipse.

Note that for $[\alpha]^{-1} = [C_{ij}]$, the eigenvectors are the same and the eigenvalues are related by $\lambda_i = \frac{1}{\lambda_i^C}$ where λ_i^C are the eigenvalues of the covariance matrix, $[C_{ij}]$. The principle

dimensions out from the centre of the ellipse are therefore: $\sqrt{\frac{\Delta\chi^2}{\lambda_1}}$ and $\sqrt{\frac{\Delta\chi^2}{\lambda_2}}$ or $\sqrt{\Delta\chi^2\lambda^C_1}$ and $\sqrt{\Delta\chi^2\lambda^C_2}$

A useful result is for the confidence region (CR) projected into one dimension to give a confidence interval (CI). This is given by Press as $\delta a_0 = \sqrt{\Delta\chi^2}\sqrt{C_{00}}$.

As $\Delta\chi^2$ is distributed as χ^2, its value can be established given the number of degrees of freedom (M) and the required confidence level. These are conventionally tabulated for σ=1, 2 etc. However any CI could be calculated, using the incomplete gamma function, $\mathrm{P}\left(\chi^2 \middle| M\right) = \frac{1}{\Gamma(M/2)}\int_0^{\Delta\chi^2/2} e^{-t}t^{M/2-1}dt$.

MONTY was tested successful on a linear regression, for which the CIs are already known. More general transport models were then added, starting with the simple AD model, and later more general 1D LT solutions to the AD-DP equation.

A number of variant codes to the basic MONTY code have been produced:

- MONTY-BOOT, uses a bootstrap of data. Note that this needs to find the mean of each parameter fits of all the distributions before the CIs either side of these are calculated. Note too that the CR must be determined numerically: the bisection method is used to adjust the size of the CR and the number insider is counted until the required percentage of points lie within the boundary.
- MONTY-LOG finds the parameter minima using logged parameters, $\log_{10}(\underline{a})$ and then recovers the parameters at the end as $10^{\underline{a}}$. This is to improve the conditioning of the shape of the error-space minimum (note PEST (Doherty, 1994)offers this as an option). However for the problems dealt with in the thesis this change made no effect on the distribution of optimised parameters (but might conceivably have speeded up each optimisation).
- MONTY-LOGN applies a log-normal error model, instead of a Normal error model.

A1.6 Ancillary codes

Further tools have been developed in support of the thesis. These are briefly summarised below:

A1.6.1 DOUBLET (Excel)

The analytical solution of Hoopes and Harleman (1965) for transport of a step input of tracer through a two-well 'doublet' system has been coded in Excel. A differential is used to give the step response.

A1.6.2 HYPERCONV adaptations (C++)

Several adapted versions of the original HYPERCONV code by Peter Robinson (Quintessa Ltd) have been produced:

HYPERCONV-F2FHIST, solves flow in system with planar boundaries (F2F) and finds a histogram of active and flowing nodes in bins that are parallel to the input and output faces and heads in each bin.

HYPERCONV-CYLHIST, solves flow in system with cylindrical boundaries and finds a histogram of active and flowing nodes in cylindrical bins that are wrapped around the inner tunnel and heads in each bin. .

HYPERCONV-SPHHIST, finds histograms of active and flowing nodes in spherical bins that are wrapped around the inner spherical boundary and heads in each bin.

HYPERCONV-PACKER, finds the potentially active nodes within a sphere (say 10 m in radius) and simulates a packer test to each node in turn (as if there was abstraction only from each discrete point). Permeabilities and flowing networks are computed. Unused in this thesis.

A1.6.3 INTERSECT (VISUAL BASIC)

Code to find the probabilities that nodes in HYPERCONV have sub-channels (bonds) attached to them (in 3D).

A1.6.4 ERROR-SPAC (FORTRAN90-MATLAB)

ERROR-SPAC is a code that plots error between data and a model (or synthetic data and a model). There is a wide choice of 'error': r^2, E_N, χ^2, SSE, MSE, RMSE, ME, MAE. FORTRAN firstly evaluates the difference between the data and the model over a grid (specified range and intensity of sampling). Then, FORTRAN uses MEX-commands to call the MATLAB® engine and plot the error. There is a choice of contours, surfaces or slices through 3D plots.

A1.6.5 VAR-VEL (FORTRAN90)

The variable flow solution to the ADE, given by Warrick et al. (1972) has been coded. This is very useful for testing the sensitivity of the ADE solution to perturbations in flow rate and therefore whether it is reasonable to assume constant flow.

Differentiating Equation A(2.3) w.r.t. z gives: $C^r = C_0^r - V\frac{\partial}{\partial z}\int_0^t C^f(z,t')dt'$ A(2.5)

Differentiating Equation A(2.4) w.r.t. z or Equation A(2.5) w.r.t. t, gives

$$\frac{\partial C^r}{\partial t} + V\frac{\partial C^f}{\partial z} = 0 \qquad \text{A(2.6)}$$

Taking LTs, Equations A(2.4) and A(2.5) become

$$\overline{C^f} = \frac{C^{f0}}{s} - \frac{1}{V}s\int_0^z \bar{C}^r dz' \qquad \text{A(2.7)}$$

and

$$\overline{C^r} = \frac{C^{r0}}{s} - V\frac{d}{dz}\frac{\bar{C}^f}{s} \qquad \text{A(2.8)}$$

In the absence of any solute initially in the system, Equation A(2.8) becomes:

$$\overline{C^r} = -\frac{V}{s}\frac{d}{dz}\bar{C}^f \qquad \text{A(2.9)}$$

B: Equating fluxes:

The flux of C^f across area A (in Figure A2.3) is VC^f A(2.10)

The flux of C^r across area A: $VC^r - D\frac{\partial C^r}{\partial z}$ A(2.11)

Assuming $\alpha V >> D_a$, so $D = \alpha V$ and combining Equations A(2.10) and A(2.11), gives

$$C^f = C^r - \alpha\frac{\partial C^r}{\partial z} \qquad \text{A(2.12)}$$

Note that the combination of Equation A(2.12) and Equation A(2.6) yields the basic Advection Dispersion Equation in C^r: i.e. $\frac{\partial C^r}{\partial t} = \alpha V\frac{\partial^2 C^r}{\partial z^2} - V\frac{\partial C^r}{\partial z}$.

If $C^r >> \alpha\frac{\partial C^r}{\partial z}$ then advection dominates over transport by dispersion, $C^f \approx C^r$, and the distinction between concentration types for all z is unimportant. Alternatively where locally $\frac{\partial C^r}{\partial z} \approx 0$, then also $C^f \approx C^r$. This last condition might be applicable to a column outlet

boundary where there is little possibility for back-dispersion from the collector (for example if the fluid drips from the column).

By substituting Equation (2.12) into $\frac{\partial C^r}{\partial t} = \alpha V \frac{\partial^2 C^r}{\partial z^2} - V \frac{\partial C^r}{\partial z}$ it can be shown that the ADE is also satisfied by C^f :

$$\frac{\partial C^f}{\partial t} = \alpha V \frac{\partial^2 C^f}{\partial z^2} - V \frac{\partial C^f}{\partial z} \qquad \text{A(2.13)}$$

In order to transform $\overline{C^f}$ and $\overline{C}^r$ in Laplace space, the following sections use Equation A(2.9), $\overline{C^r} = -\frac{V}{s}\frac{d\overline{C}^f}{dz}$ (assuming no initial resident concentration in the system) and the LT of Equation (2.12), $\overline{C^f} = \overline{C^r} - \alpha \frac{d\overline{C^r}}{dz}$.

There is a wide range of published solutions to the ADE depending upon the assumed boundary or initial conditions.[1] The most important 1D solutions are discussed in the following section, taking into account the definition of C.

The well-known boundary condition 'types' can now be specified a little more subtly in terms of the following four conditions:

(a) A 'Type 1' constant C^r boundary (otherwise known as a Dirichlet boundary).

(b) A 'Type 3' boundary is described by constant $C^f = C^r - \alpha \frac{\partial C^r}{\partial z}$ (otherwise known as a Cauchy boundary)

(c) A 'Type 1' constant C^f boundary (which is effectively the same as condition (b))

(d) A 'Type 3' boundary described by constant $C = C^f - \alpha \frac{\partial C^f}{\partial z}$ has no obvious physical meaning.

This of course boils down to only two practical conditions. Note that for condition (a), the advective component of transport (VC^r) is constant, but the total solute flux is not specified.

[1] van Genuchten and Alves (1982) gives comprehensive solutions for continuous input. For infinite downstream boundaries van Genuchten (1999) and Toride et al. (1993) give comprehensive solutions. Jury and Roth (1990) gives an introduction to Fourier and LT solution methods and presents a number of solutions.

For Type 1' constant C^f, the total solute flux (VC^f) is constant, but the advective component is not specified.

There are four permutations possible for inlet boundary and type of concentration, which are applicable to different input and measurement scenarios (Table A2.1).

Table A2.1: Permutations of inlet and measured concentrations for column experiments

Permutation	Inlet BC	Measured concentration type
1	Type 1 in C^r	C^r (internal)
2	Type 3 in C^r or Type 1 in C^f	C^f (outlet)
3	Type 3 in C^r or Type 1 in C^f	C^r (internal)
4	Type 1 in C^r	C^f (outlet)

Some care is therefore required in correctly assigning the analytical solution to different situations, if the method of concentration measurement is to be correctly taken into account.

A2.3 Impulse response

A2.3.1 Infinite downstream boundary

The unit impulse responses for $C^f = (m_0/V)\delta(t)$ and $C^r = (m_0/V)\delta(t)$ are given in Table A2.2. It is clear that the definition of concentration and the assumed inlet boundary make a considerable difference to the functional form of the response.

Table A2.2: Impulse responses to ADE for different boundary conditions and concentration types (solutions achieved in Derive).

Impulse input	Laplace solution	Time solution	Equation number
$C^f(0,t) = (m_0/V)\delta(t)$	$\overline{C^r}(z,s) = \dfrac{2(m_0/V)}{\left(1+\sqrt{1+\dfrac{4s\alpha}{v}}\right)} e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{v}}\right)}$	$C^r(z,t) = \dfrac{V(m_0/V)}{\sqrt{\pi Dt}} e^{-\frac{(z-Vt)^2}{4Dt}} - \dfrac{V^2(m_0/V)}{2D} e^{Vz/D}\operatorname{erfc}\left(\dfrac{z+Vt}{\sqrt{4Dt}}\right)$ *	A(2.14), A(2.15)
$C^f(0,t) = (m_0/V)\delta(t)$	$\overline{C^f}(z,s) = (m_0/V) e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$C^f(z,t) = \dfrac{z(m_0/V)}{\sqrt{4\pi Dt^3}} e^{-\frac{(z-Vt)^2}{4Dt}}$	A(2.16), A(2.17)
$C^f(0,t) = (m_0/V)\delta(t)$	$\overline{C^r}(z,s) = (m_0/V) e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$C^r(z,t) = \dfrac{z(m_0/V)}{\sqrt{4\pi Dt^3}} e^{-\frac{(z-Vt)^2}{4Dt}}$	A(2.18), A(2.19)
$C^f(0,t) = (m_0/V)\delta(t)$	$\overline{C^f}(z,s) = (m_0/V) e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$ $-\dfrac{1}{2}(m_0/V)\left(1-\sqrt{1+\dfrac{4s\alpha}{v}}\right) e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$C^f(z,t) = (m_0/V)\left(\dfrac{z+2tVz-2\alpha Vt}{4\sqrt{\pi\alpha t^5 V^3}}\right) e^{-\frac{(z-Vt)^2}{4Dt}}$	A(2.20), A(2.21)

* Note that Equation 3.12 in Jury and Roth, 1990, p.48 is identical to this except for an erroneous θ.

A2.3.2 Finite downstream boundary

Unit impulse response solutions to the ADE for finite boundaries have been found in Laplace space.[2] For a 'Type 3' C^r inlet boundary condition (i.e. Type 1, $C^f = (m_0/V)\delta(t)$) and an outlet boundary condition of $\left.\frac{dC^r}{dx}\right|_{x=L} = 0$:

$$\bar{C}^f(L) = (m_0/V)\frac{4\varsigma}{(1+\varsigma)^2 - (1-\varsigma)^2 \exp\left(-\frac{V_m L\varsigma}{D_m}\right)} \underbrace{\exp\left(\frac{V_m L}{2D_m}(1-\varsigma)\right)}_{Inf\_BC} \qquad \text{A(2.22)}$$

where $\varsigma = \sqrt{1 + \frac{4sD_m}{V_m^2}}$.

The solution in terms of resident concentration is:

$$\bar{C}^r(L) = 2(m_0/V)\left[\frac{-(1-\varsigma)\exp\left(\frac{VL(1-\varsigma)}{2D}\right) + (1+\varsigma)\exp\left(\frac{VL(1+\varsigma)}{2D}\right)}{(1+\varsigma)^2 - (1-\varsigma)^2 \exp\left(-\frac{VL\varsigma}{D}\right)}\right] \qquad \text{A(2.23)}$$

For $C^r = \delta(t)$ and $\left.\frac{dC^r}{dz}\right|_{z=L} = 0$:

$$\bar{C}^r(L) = (m_0/V)\frac{2\varsigma \exp\left(\frac{V_m L}{2D_m}(1-\varsigma)\right)}{\left(1+\varsigma-(1-\varsigma)\exp\left(-\frac{VL\varsigma}{D}\right)\right)} \qquad \text{A(2.24)}$$

The solution for C^f is established using $C^f(L) = C^r(L) - \alpha\left.\frac{dC^r}{dz}\right|_{z=L}$, which yields a rather complicated expression (not shown here).

[2] Working is not shown, for brevity. All solutions have been confirmed by substitution back into the ADE.

A2.4 Step response

A2.4.1 Infinite downstream boundary

The semi-infinite step response is given in Table A2.3. This is the time integral of the impulse response.

Table A2.3: Semi-infinite step response.

Input	Laplace	Time	Author	Equation number
$C^f = C_0 H(t)$	$\frac{\overline{C^r}(z,s)}{C_0} = -\frac{1}{s}\frac{2}{\left(1+\sqrt{1+\frac{4s\alpha}{V}}\right)}e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$\frac{C^r(z,t)}{C_0} = \frac{1}{2}\operatorname{erfc}\left(\frac{z-Vt}{\sqrt{4Dt}}\right)+\left(\frac{V^2 z}{\pi D}\right)^{1/2} e^{\frac{-(z-Vt)^2}{4Dt}}$ $-\frac{1}{2}\left[1+\frac{Vz}{D}+\frac{V^2 t}{D}\right]e^{\frac{Vz}{D}}\operatorname{erfc}\left(\frac{z+Vt}{\sqrt{4Dt}}\right)$	Gershon and Nir (1969) [3]	A(2.25) A(2.26)
$C^f = C_0 H(t)$	$\frac{\overline{C^f}(z,s)}{C_0} = \frac{1}{s}e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$\frac{C^f}{C_0} = \frac{1}{2}\operatorname{erfc}\left(\frac{z-Vt}{\sqrt{4Dt}}\right)+\frac{1}{2}e^{(Vz/D)}\operatorname{erfc}\left(\frac{z+Vt}{\sqrt{4Dt}}\right)$	Lapidus and Amundson (1952), Ogata and Banks (1961)	A(2.27), A(2.28)
$C^r = C_0 H(t)$	$\frac{\overline{C^r}(z,s)}{C_0} = \frac{1}{s}e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$\frac{C^r}{C_0} = \frac{1}{2}\operatorname{erfc}\left(\frac{z-Vt}{\sqrt{4Dt}}\right)+\frac{1}{2}e^{(Vz/D)}\operatorname{erfc}\left(\frac{z+Vt}{\sqrt{4Dt}}\right)$	Lapidus and Amundson (1952), Ogata and Banks (1961)	A(2.29), A(2.30)
$C^r = C_0 H(t)$	$\frac{\overline{C^f}(z,s)}{C_0} = \frac{1}{s}e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$ $-\frac{1}{2s}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)e^{\frac{z}{2\alpha}\left(1-\sqrt{1+\frac{4s\alpha}{V}}\right)}$	$\frac{C^f}{C_0} = \int_0^t \left(\frac{z+2t'Vz-2Dt'}{4\sqrt{\pi Dt'^5 V^2}}\right)e^{-\frac{(z-Vt')^2}{4Dt'}}dt'$ †	Here	A(2.31), A(2.32)

†Expressed as an integral because, although the full solution is analytically obtainable, it is long and cumbersome

[3] Appears in most standard texts (but is incorrectly cited in Fetter (1999))

A2.4.2 Finite boundary step solution summary

The LT step solutions where there are finite downstream boundary conditions are $1/s$ times the impulse response to the same boundary conditions. The inverse LT gives the time domain solutions.

For $C^r(0,t) = C_0 \mathrm{H}(t)$ and $\left.\frac{dC^r}{dz}\right|_{z=L} = 0$ the time-domain solution is given by Cleary and Adrian, 1973:

$$\frac{C^r(z)}{C_0} = 1 - \sum_{m-1}^{\infty} \frac{2\beta_m \sin\left(\frac{\beta_m z}{L}\right) \exp\left(\frac{Vz}{2D} - \frac{V^2 t}{4D} - \frac{\beta_m{}^2 Dt}{L^2}\right)}{\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2 + \frac{VL}{2D}} \qquad \text{A(2.33)}$$

where the eigenvalues β_m are the positive roots of: $\beta_m \cot(\beta_m) + \frac{VL}{2D} = 0$.

For $C^f(0,t) = C_0 H(t)$ and $\left.\frac{dC^r}{dz}\right|_{z=L} = 0$ the solution is given by Brenner, 1962:

$$\frac{C^r(z)}{C_0} = 1 - \sum_{m-1}^{\infty} \frac{\frac{2VL}{D}\beta_m \left[\beta_m \cos\left(\frac{\beta_m z}{L}\right) + \frac{VL}{2D}\sin\left(\frac{\beta_m z}{L}\right)\right] \exp\left(\frac{Vz}{2D} - \frac{V^2 t}{4DR} - \frac{\beta_m{}^2 Dt}{L^2 R}\right)}{\left(\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2 + \frac{VL}{D}\right)\left(\beta_m{}^2 + \left(\frac{VL}{2D}\right)^2\right)} \qquad \text{A (2.34)}$$

where the eigenvalues β_m are the positive roots of: $\beta_m \cot(\beta_m) - \frac{\beta_m{}^2 D}{VL} + \frac{VL}{4D} = 0$

The appropriate solutions for different concentration types for finite boundaries are given in Table A2.4. The special (and probably most useful) result is that the volume averaged output is the same as the flux-averaged output, since $\left.\frac{dC^r}{dx}\right|_{z=L} = 0$.

be a poor approximation. However, it is not unambiguously clear what finite boundary condition is correct to apply.

Parlange et al., 1992 showed that this ambiguity means that the condition for C^r at $z = L$ is bounded by two extreme conditions: $\left.\frac{\partial C^r}{\partial z}\right|_L = 0$ and $\left.\frac{\partial C^r}{\partial z}\right|_\infty = 0$.

A2.6 Effect of boundary conditions on late-time gradients

A2.6.1 Spatial input

The response for the ADE to the unit spatial input of $C^r(z,0) = m_0\delta(z)$, with infinite upstream and downstream boundary conditions is: $C^r = \frac{m_0}{\sqrt{4\pi\alpha Vt}} e^{\frac{-(z-Vt)^2}{4\alpha Vt}}$.

Taking logs of both sides and differentiating w.r.t. t gives,

$$\frac{\partial}{\partial t}\left[\ln(C^r)\right] = \left(\frac{z^2}{4\alpha V}\right)\frac{1}{t^2} - \left(\frac{1}{2}\right)\frac{1}{t} - \frac{V}{4\alpha}.$$

It is useful to compare this to the semi-infinite response. The lin-log gradient, $-V/4\alpha$, is the same for both solutions. However, remarkably, the gradient for the log-log gradient is -1/2 for the infinite boundary condition, rather than -3/2 for the semi-infinite boundary. Boundary conditions therefore have a significant effect on the late-time behaviour of the AD model.

A2.6.2 Late time gradient for flushing system

The flushing response for a semi infinite system with a unit constant initial concentration, G_F, is simply $G_F = 1 - G_H$, where G_H is the Ogata Banks step response. This is plotted in Figure A2.6. It can be seen that the gradient of $\ln(C^r)$ against time converges with the gradient of the impulse response.

$$C^r = \frac{1}{2}\left\{\text{erfc}\left(\frac{z-Vt}{\sqrt{4\alpha vt}}\right)+\frac{1}{2}e^{\frac{z}{\alpha}}\text{erfc}\left(\frac{z+Vt}{\sqrt{4\alpha Vt}}\right)\right\}$$
$$-e^{-\lambda t}\left\{\frac{1}{2}e^{\frac{(V-y)z}{2\alpha V}}\text{erfc}\left(\frac{a-yt}{\sqrt{4\alpha Vt}}\right)+\frac{1}{2}e^{\frac{(V+y)a}{2\alpha V}}\text{erfc}\left(\frac{a+yt}{\sqrt{4\alpha Vt}}\right)\right\} \quad \text{A(2.35)}$$

where, $y = V\sqrt{1-4\lambda/V}$

This yields a very complicated analytical solution when manipulated in order to find the late-time log-log and lin-log gradients (and is therefore excluded here).

The gradients of the ADE inline with a mixer were therefore calculated by numerically inverting the convolution of the mixer function and the solution to the semi-infinite ADE and then plotting the results. This showed that the lin-log gradient still tends to $-V/4\alpha$, although the time at which this is reached is delayed when there is a substantial mixing cell. In the instance of high dispersion, where a period of log-log -3/2 gradient would be generated in the absence of a mixer, a substantial mixer can completely eliminate the log-log linear period.

A2.6.4 Summary of Gradients

Table A2.5 gives a summary of the different late-time gradients predicted for the ADE with different combinations of inlet boundary conditions and outlet concentrations.

Table A2.5: $\frac{\partial}{\partial t}\ln(C)$**, for the semi-infinite solution to the ADE with different boundary conditions.**

Inlet BC	Outlet concentration: Resident	Outlet concentration: Flux
$C^f = \delta(t)$	$Z^* + \frac{z^2}{4\alpha Vt^2} - \frac{V}{4\alpha}$	$\left(\frac{z^2}{4\alpha V}\right)\frac{1}{t^2} - \left(\frac{3}{2}\right)\frac{1}{t} - \frac{V}{4\alpha}$
$C^r = \delta(t)$	$\left(\frac{z^2}{4\alpha V}\right)\frac{1}{t^2} - \left(\frac{3}{2}\right)\frac{1}{t} - \frac{V}{4\alpha}$	$-\frac{2\alpha V}{(2tV+1)[z(2tV+1)-2\alpha tV]} + \frac{z^2}{4\alpha Vt^2}$ $+\frac{2V}{(2tV+1)} - \frac{V}{4\alpha} - \frac{5}{2t}$

where Z^* is $\frac{\partial}{\partial t}\ln\left[\frac{V\left(V\sqrt{\pi Dt}e^{-\frac{(z-Vt)^2}{4Dt}}\operatorname{erfc}\left(\frac{(z+Vt)}{2\sqrt{Dt}}\right)+2D\right)}{2\sqrt{D^3t}}\right]$

The lin-log gradients in Table A2.5 are all $-V/4\alpha$ and log-log gradients are given in Table A2.6.

Table A2.6: Log-log Gradients

Input	Output	
$\frac{\partial}{\partial t}\ln(C)$	**Resident**	**Flux**
$C^f = \delta(t)$	None	-3/2
$C^r = \delta(t)$	-3/2	-5/2

A2.7 Effect of boundary conditions on moments

Table A2.7 gives the first moments and Table A2.8 gives the second moments predicted for the ADE with different combinations of inlet boundary conditions and outlet concentrations. The variety of solutions means that care needs to be taken over the boundary conditions and concentration type if using the method of moments to interpret parameter values.

Table A2.7: First temporal moments for the AD model.

Boundary conditions		Output at $z = L$	
Upstream	**Downstream**	**Resident**	**Flux**
	$C^f = \delta(t)$ $\left.\frac{\partial C^r}{\partial z}\right\|_\infty = 0$	$\frac{L}{V}+\frac{\alpha}{V}$	$\frac{L}{V}$
$C^r = \delta(t)$	$\left.\frac{\partial C^r}{\partial z}\right\|_\infty = 0$	$\frac{L}{V}$	$\frac{L}{V}-\frac{\alpha}{V}$
$C^f = \delta(t)$	$\left.\frac{\partial C^r}{\partial z}\right\|_L = 0$	$\frac{L}{V}$	$\frac{L}{V}$
$C^r = \delta(t)$	$\left.\frac{\partial C^r}{\partial z}\right\|_L = 0$	$\frac{L}{V}-\frac{\alpha}{V}+\frac{\alpha}{V}e^{-L/\alpha}$	Not found[5]

[5] Derive was unable to find a solution.

Note that the solution for $C^r = \delta(t)$, $\left.\frac{\partial C^r}{\partial z}\right|_{\infty} = 0$ and flux output only makes physical sense if $\alpha < L$. There is also a constraint on $C^r = \delta(t)$, $\left.\frac{\partial C^r}{\partial z}\right|_{L} = 0$ and resident output.

Table A2.8: Second centred temporal moments for the AD model

	Resident	**Flux**
$C^f = \delta(t)$ $\left.\frac{\partial C^r}{\partial z}\right\|_{\infty} = 0$	$\frac{\alpha}{V^2}(2L+3\alpha)$	$\frac{2\alpha L}{V^2}$
$C^r = \delta(t)$ $\left.\frac{\partial C^r}{\partial z}\right\|_{\infty} = 0$	$\frac{2\alpha L}{V^2}$	$\frac{\alpha}{V^2}(2L-3\alpha)$
$C^f = \delta(t)$ $\left.\frac{\partial C^r}{\partial z}\right\|_{L} = 0$	$\frac{2\alpha^2}{V^2}e^{-LV/D} - \frac{2\alpha(\alpha - L)}{V^2}$	$\frac{2\alpha^2}{V^2}e^{-LV/D} - \frac{2\alpha(\alpha - L)}{V^2}$
$C^r = \delta(t)$ $\left.\frac{\partial C^r}{\partial z}\right\|_{L} = 0$	$\frac{4\alpha(\alpha+L)}{V^2}e^{-L/\alpha} + \frac{\alpha^2}{V^2}e^{2L/\alpha} - \frac{\alpha(5\alpha-2L)}{V^2}$	Not found

A3 APPENDIX 3 – HyperConv probabilities

A3.1 Introduction

The channel network created by HyperConv (written by Peter Robinson) is based on a simple stochastic generator for allocating 'sub-channels' on a cubic grid of nodes. Sub-channels are defined as connections ('bonds') across the gaps between adjacent nodes and 'channels' are defined as being one or more sub-channels in a line (see Figure 13.3 in the main text). At either end of the channel must either be a boundary or a gap (a lack of sub-channels) . HyperConv achieves a complete network by working systematically along the x, y and z directions allocating 'sub-channels' according to a simple algorithm:

- The first sub-channel is generated with a probability, P_{ON} .
- The next sub-channel is generated with probability P_A if there was a first sub-channel.
- If there was no first sub-channel, then a sub-channel between the next nodes is generated with probability P_N .
- The application of P_N if there was no previous sub-channel and P_A , if there was a previous sub-channel, is repeated sequentially across all nodes in the line.

The statistics of channel lengths and of the gaps between channels that arise due to this generating algorithm are considered here. These were established by this author and systematically gathered into tables by John Barker.

A3.2 Basic probability relationships

Following a given first sub-channel of a particular channel, the probability that there is another sub-channel along the generating direction is P_A. Hence, the probability of a channel having a particular length, L , (in units of number of sub-channels) is given by

$$P\{L=1\}=(1-P_A)$$

$$P\{L=2\}=P_A(1-P_A)$$

$$P\{L=3\}=P_A^2(1-P_A)$$

$$P\{L=N\}=P_A^{N-1}(1-P_A) \qquad \text{A(3.1)}$$

These arise from setting the probability of the first sub-channel existing as one and then there being $N-1$ sub-channels with probability P_A and then finally there must be a 'gap' defining the end of the channel, with probability $1-P_A$.

By inspection, the mode (maximum) occurs at $L=1$, with magnitude $(1-P_A)$.

The mean, $\overline{L}$, is given by

$$\overline{L}=\frac{\sum_{N=1}^{\infty}NP_A^{N-1}(1-P_A)}{\sum_{N=1}^{\infty}P_A^{N-1}(1-P_A)}=\frac{\sum_{N=1}^{\infty}NP_A^{N-1}}{\sum_{N=1}^{\infty}P_A^{N-1}}=\frac{1/(1-P_A)^2}{1/(1-P_A)}=\frac{1}{(1-P_A)} \qquad \text{A(3.2)}$$

The variance, σ_L^2, is obtained as

$$\sigma_L^2=\sum_{N=1}^{\infty}\left[N-\frac{1}{(1-P_A)}\right]^2P_A^{N-1}(1-P_A)=\frac{P_A}{(1-P_A)^2} \qquad \text{A(3.3)}$$

Following the first gap (lack of sub-channel), the next gap has a probability of P_N. Hence, the probability of a line of gaps having a particular length G (derived in the same was as the channel probabilities) is given by

$$\begin{aligned}P\{G=1\}&=P_N\\P\{G=2\}&=P_N(1-P_N)\\P\{G=3\}&=P_N(1-P_N)^2\\P\{G=N\}&=P_N(1-P_N)^{N-1}\end{aligned} \qquad \text{A(3.4)}$$

The mode (maximum) occurs at $G=1$, with magnitude P_N.

The mean, $\overline{G}$, is obtained as follows:

$$\overline{G}=\frac{\sum_{N=1}^{\infty}NP_N(1-P_N)^{N-1}}{\sum_{N=1}^{\infty}P_N(1-P_N)^{N-1}}=\frac{\sum_{N=1}^{\infty}N(1-P_N)^{N-1}}{\sum_{N=1}^{\infty}(1-P_N)^{N-1}}=\frac{1/P_N{}^2}{1/P_N}=\frac{1}{P_N} \qquad \text{A(3.5)}$$

The variance, σ_G^2, is therefore

$$\sigma_G^2=\sum_{N=1}^{\infty}\left[N-\frac{1}{P_N}\right]^2P_N(1-P_N)^{N-1}=\frac{1-P_N}{P_N^2} \qquad \text{A(3.6)}$$

The probability that any given sub-channel belongs to a channel of length N is proportional to the channel length, N, and proportional to the probability of that length, as given by A(3.1). Therefore it is proportional to

$$NP_A^{N-1}(1-P_A) \qquad \text{A(3.7)}$$

To obtain the probability, we normalize by the sum of this over all N:

$$\sum_{N=1}^{\infty} NP_A^{N-1}(1-P_A) \qquad \text{A(3.8)}$$

So the probability that a chosen sub-channel belongs to a channel of length N is:

$$P\{\text{sub-channel belongs to channel of length} = N\}$$
$$= \frac{NP_A^{N-1}(1-P_A)}{\sum_{N=1}^{\infty} NP_A^{N-1}(1-P_A)} = NP_A^{N-1}(1-P_A)^2 \qquad \text{A(3.9)}$$

Using this result, the mean length of channel that a sub-channel might belong to is given by

$$\overline{C} = \sum_{N=1}^{\infty} NNP_A^{N-1}(1-P_A)^2 = \frac{1+P_A}{1-P_A} \qquad \text{A(3.10)}$$

Similarly, the variance, σ_C^2, is given by

$$\sigma_c^2 = \sum_{N=1}^{\infty}\left[N - \frac{(1+P_A)}{(1-P_A)}\right]^2 NP_A^{N-1}(1-P_A)^2 = \frac{2P_A}{(1-P_A)^2} \qquad \text{A(3.11)}$$

The mode is obtained by solving $\frac{d}{dN}\left[NP_A^{N-1}(1-P_A)^2\right] = 0$, which gives

$$-\frac{1}{\ln(P_A)} \qquad \text{A(3.12)}$$

A3.3 Probability that a node has sub-channel attached

Denote the probability that a node has at least one sub-channel connected to it along a line of generation as $\hat{P}_{ON}$. This is simply calculated by first establishing the probability that there is no sub-channel to the side of the node:

$P\{\text{No sub-channel upstream and down-stream of node}\} = (1-P_{ON})(1-P_N)$.

Therefore the probability that one or more neighbouring sub-channels exists is simply:

$$\hat{P}_{ON} = 1-(1-P_{ON})(1-P_N) \qquad \text{A(3.13)}$$

This can be rearranged by substituting for $P_N = \frac{P_{ON}(1-P_A)}{1-P_{ON}}$, to

$$\hat{P}_{ON} = P_{ON}(2-P_A) \qquad \text{A(3.14)}$$

When the three-dimensionality of the system is taken into account this gives a 3D probability that a node has at least one sub-channel connected, $\hat{P}_{ON}^{3D}$. Following similar logic to the 1D case, the probability can be found from:

$P\{\text{No sub-channel upstream and down-stream of node in all directions}\} = (1-P_{ON})^3(1-P_N)^3$

giving

$$\hat{P}_{ON}^{3D} = 1-(1-P_{ON})^3(1-P_N)^3 \qquad \text{A(3.15)}$$

Alternatively $\hat{P}_{ON}^{3D}$ can be expressed as a function of $\hat{P}_{ON}$:

$$\hat{P}_{ON}^{3D} = \hat{P}_{ON}^{1D}\ (3-3\ \hat{P}_{ON}^{1D} + \hat{P}_{ON}^{1D}\ ) \qquad \text{A(3.16)}$$